An Asymptotically Optimal Algorithm for the Max k-Armed Bandit Problem

Matthew J. Streeter **Stephen F. Smith**

February 27, 2006
CMU-CS-06-110

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Abstract

We present an asymptotically optimal algorithm for the *max* variant of the k-armed bandit problem. Given a set of k slot machines, each yielding payoff from a fixed (but unknown) distribution, we wish to allocate trials to the machines so as to maximize the expected maximum payoff received over a series of n trials. Subject to certain distributional assumptions, we show that $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ trials are sufficient to identify, with probability at least $1-\delta$, a machine whose expected maximum payoff is within ϵ of optimal. This result leads to a strategy for solving the problem that is asymptotically optimal in the following sense: the gap between the expected maximum payoff obtained by using our strategy for n trials and that obtained by pulling the single best arm for all n trials approaches zero as $n \to \infty$.

Keywords: multi-armed bandit problem, PAC bounds

1 Introduction

In the k-armed bandit problem one is faced with a set of k slot machines, each having an arm that, when pulled, yields a payoff from a fixed (but unknown) distribution. The goal is to allocate trials to the arms so as to maximize the expected cumulative payoff obtained over a series of n trials. Solving the problem entails striking a balance between exploration (determining which arm yields the highest mean payoff) and exploitation (repeatedly pulling this arm).

In the max k-armed bandit problem, the goal is to maximize the expected *maximum* (rather than cumulative) payoff. This version of the problem arises in practice when tackling combinatorial optimization problems for which a number of randomized search heuristics exist: given k heuristics, each yielding a stochastic outcome when applied to some particular problem instance, we wish to allocate trials to the heuristics so as to maximize the maximum payoff (e.g., the maximum number of clauses satisfied by any sampled variable assignment, the minimum makespan of any sampled schedule). Cicirello and Smith [3] show that a max k-armed bandit approach yields good performance on the resource-constrained project scheduling problem with maximum time lags (RCPSP/max).

1.1 Summary of Results

We consider a restricted version of the max k-armed bandit problem in which each arm yields payoff drawn from a *generalized extreme value (GEV) distribution* (defined in §2). This paper presents the first provably asymptotically optimal algorithm for this problem.

Roughly speaking, the reason assuming a GEV distribution is the Extremal Types Theorem (stated in §2), which states that the distribution of the sample maximum of n independent identically distributed random variables approaches a GEV distribution as $n \rightarrow \infty$. A more formal justification is given in §3. For reasons that will become clear, the nature of our results depend on the shape parameter (ξ) of the GEV distribution. Assuming all arms have $\xi \leq 0$, our results can be summarized as follows.

- Let a be an arm that yields payoff drawn from a GEV distribution with unknown parameters; let M_n denote the maximum payoff obtained after pulling a n times; and let $m_n = \mathbb{E}[M_n]$. We provide an algorithm that, after pulling the arm $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ times, produces an estimate $\bar{m}_n$ of m_n with the property that $\mathbb{P}\left[|\bar{m}_n - m_n| < \epsilon\right] \geq 1 - \delta$.

- Let $a_1, a_2, \ldots, a_k$ be k arms, each yielding payoff from (distinct) GEV distributions with unknown parameters. Let m_n^i denote the expected maximum payoff obtained by pulling the i^{th} arm n times, and let $m_n^* = \max_{1 \leq i \leq k} m_n^i$. We provide an algorithm that, when run for n pulls, obtains expected maximum payoff $m_n^* - o(1)$.

Our results for the case $\xi > 0$ are similar, except that our estimates and expected maximum payoffs come within arbitrarily small *factors* (rather than absolute distances) of optimality. Specifically, our estimates have the property that $\mathbb{P}\left[\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < 1 + \epsilon\right] \geq 1 - \delta$ for constant α_1 independent of n, while the expected maximum payoff obtained by using our algorithm for n pulls is $m_n^*(1 - o(1))$.

1.2 Related Work

The classical k-armed bandit problem was first studied by Robbins [7] and has since been the subject of numerous papers; see Berry and Fristedt [1] and Kaelbling [6] for overviews. In a paper similar in spirit to ours, Fong [5] showed that an initial exploration phase consisting of $O(\frac{k}{\epsilon^2}\ln(\frac{k}{\delta}))$ pulls is sufficient to identify, with probability at least $1-\delta$, an arm whose mean payoff is within ϵ of optimal. Theorem 2 of this paper proves a bound similar to Fong's on the number of pulls needed to identify an arm whose expected *maximum* payoff (over a series of n trials) is near-optimal.

The max variant of the k-armed bandit problem was first studied by Cicirello and Smith [2, 3], who successfully used a heuristic for the max k-armed bandit problem to select among priority rules for the RCPSP/max. The design of Cicirello and Smith's heuristic is motivated by an analysis of the special case in which each arm's payoff distribution is a GEV distribution with shape parameter $\xi = 0$, but they do not rigorously analyze the heuristic's behavior. Our paper is more theoretical and less empirical: on the one hand we do not perform experiments on any practical combinatorial problem, but on the other hand we provide stronger performance guarantees under weaker distributional assumptions.

1.3 Notation

For an arbitrary cumulative distribution function G, let the random variable M_n^G be defined by

$$M_n^G = \max\{Z_1, Z_2, \ldots, Z_n\}$$

where $Z_1, Z_2, \ldots, Z_n$ are independent random variables, each having distribution G. Let

$$m_n^G = \mathbb{E}[M_n^G]\ .$$

2 Extreme Value Theory

This section provides a self-contained overview of results in extreme value theory that are relevant to this work. Our presentation is based on the text by Coles [4].

The central result of extreme value theory is an analogue of the central limit theorem that applies to extremely rare events. Recall that the central limit theorem states that (under certain regularity conditions) the distribution of the sum of n independent, identically distributed (i.i.d) random variables converges to a normal distribution as $n \to \infty$. The extremal types theorem states that (under certain regularity conditions) the distribution of the maximum of n i.i.d random variables converges to a generalized extreme value (GEV) distribution.

Definition (GEV distribution). *A random variable Z has a* generalized extreme value distribution *if, for constants μ, $\sigma > 0$, and ξ, $\mathbb{P}[Z \leq z] = GEV_{(\mu,\sigma,\xi)}(z)$, where*

$$GEV_{(\mu,\sigma,\xi)}(z) = \exp\left(-\left(1 + \frac{\xi(z-\mu)}{\sigma}\right)^{-\frac{1}{\xi}}\right)$$

for $z \in \{z : 1+\xi(z-\mu)\sigma^{-1} > 0\}$*, and* $GEV_{(\mu,\sigma,\xi)}(z) = 1$ *otherwise. The case* $\xi = 0$ *is interpreted as the limit*

$$\lim_{\xi' \to 0} GEV_{(\mu,\sigma,\xi')}(z) = \exp\left(-\exp\left(\frac{\mu - z}{\sigma}\right)\right) .$$

The following three propositions establish properties of the GEV distribution.

Proposition 1. *Let* Z *be a random variable with* $\mathbb{P}[Z \leq z] = GEV_{(\mu,\sigma,\xi)}(z)$*. Then*

$$\mathbb{E}[Z] = \begin{cases} \mu + \frac{\sigma}{\xi}\left(\Gamma(1-\xi) - 1\right) & \text{if } \xi < 1, \xi \neq 0 \\ \mu + \sigma\gamma & \text{if } \xi = 0 \\ \infty & \text{if } \xi \geq 1 \end{cases}$$

where

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

is the complete gamma function and

$$\gamma = \lim_{n \to \infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \ln(n) \right)$$

is Euler's constant.

Proposition 2. *Let* $G = GEV_{(\mu,\sigma,\xi)}$*. Then* M_n^G *has distribution* $GEV_{(\mu',\sigma',\xi')}$*, where*

$$\mu' = \begin{cases} \mu + \frac{\sigma}{\xi}(n^\xi - 1) & \text{if } \xi \neq 0 \\ \mu + \sigma \ln(n) & \text{otherwise,} \end{cases}$$
$$\sigma' = \sigma n^\xi \text{, and}$$
$$\xi' = \xi \, .$$

Substituting the parameters of M_n^G given by Proposition 2 into Proposition 1 gives an expression for m_n^G.

Proposition 3. *Let* $G = GEV_{(\mu,\sigma,\xi)}$ *where* $\xi < 1$*. Then*

$$m_n^G = \begin{cases} \mu + \frac{\sigma}{\xi}\left(n^\xi \Gamma(1-\xi) - 1\right) & \text{if } \xi \neq 0 \\ \mu + \sigma\gamma + \sigma \ln(n) & \text{otherwise.} \end{cases}$$

It follows that

- *for* $\xi > 0$, m_n^G *is* $\Theta(n^\xi)$*;*
- *for* $\xi = 0$, m_n^G *is* $\Theta(\ln(n))$*; and*
- *for* $\xi < 0$, $m_n^G = \mu - \frac{\sigma}{\xi} - \Theta(n^\xi)$.

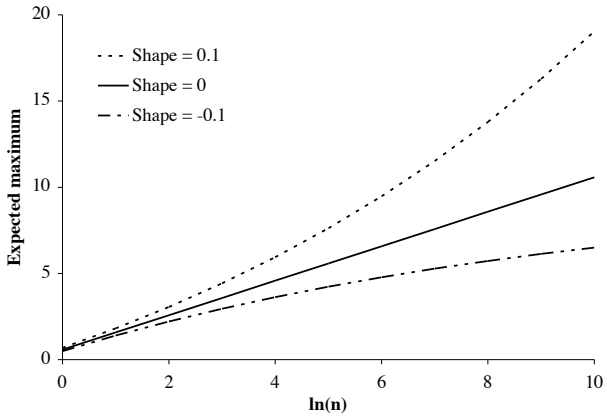

Figure 1: The effect of the shape parameter (ξ) on the expected maximum of n independent draws from a GEV distribution.

It is useful to have a visual picture of what Proposition 3 means. Figure 1 plots m_n^G as a function of n for three GEV distributions with $\mu = 0$, $\sigma = 1$, and $\xi \in \{0.1, 0, -0.1\}$.

The central result of extreme value theory is the following theorem.

The Extremal Types Theorem. *Let G be an arbitrary cumulative distribution function, and suppose there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that*

$$\lim_{n\to\infty} \mathbb{P}\left[\frac{M_n^G - b_n}{a_n} \leq z\right] = G^*(z) \tag{2.1}$$

for any continuity point z of G^*, where G^* is a not a point mass. Then there exist constants μ, $\sigma > 0$, and ξ such that $G^*(z) = GEV_{(\mu,\sigma,\xi)}(z)\ \forall z$. Furthermore,

$$\lim_{n\to\infty} \mathbb{P}\left[M_n \leq z\right] = GEV_{(\mu a_n + b_n, \sigma a_n, \xi)}(z)\ .$$

Condition (2.1) holds for a variety of distributions including the normal, lognormal, uniform, and Cauchy distributions.

3 The Max k-Armed Bandit Problem

Definition (max k-armed bandit instance). *An instance $I = (n, \mathcal{G})$ of the max k-armed bandit problem is an ordered pair whose first element is a positive integer n, and whose second element is a set $\mathcal{G} = \{G_1, G_2, \ldots, G_k\}$ of k cumulative distribution functions, each thought of as an arm on a slot machine. The i^{th} arm, when pulled, returns a realization of a random variable with distribution G_i.*

Definition (max k-armed bandit strategy). *A max k-armed bandit strategy S is an algorithm that, given an instance $I = (n, \mathcal{G})$ of the max k-armed bandit problem, performs a sequence of n*

arm-pulls. For any strategy S and integer $\ell \leq n$, we denote by $S_\ell(I)$ the expected maximum payoff obtained by running S on I for ℓ trials:

$$S_\ell(I) = \mathbb{E}\left[\max_{0 \leq j \leq \ell} p_j\right]$$

where p_j is the payoff obtained from the j^{th} pull, and we define $p_0 = 0$.

Our goal is to come up with a strategy S such that $S_n(I)$ is near-maximal.

Note that the problem is ill-posed (i.e., there is no clear criterion for preferring one strategy over another) unless we make some assumptions about the distributions G_i. We will assume that each arm $G_i = GEV_{(\mu_i,\sigma_i,\xi_i)}$ is a GEV distribution whose parameters satisfy

1. $|\mu_i| \leq \mu_u$
2. $0 < \sigma_\ell \leq \sigma_i \leq \sigma_u$
3. $\xi_\ell \leq \xi_i \leq \xi_u < \frac{1}{2}$

for known constants μ_u, σ_ℓ, σ_u, ξ_ℓ, and ξ_u.

There are two arguments for assuming that each arm is a GEV distribution. First, in practice the distribution of payoffs returned by a strong heuristic may be approximately GEV, even if the conditions of the Extremal Type Theorem are not formally satisfied [2].

A second argument runs as follows. Suppose $I = (n, \mathcal{G})$ is an instance of the max k-armed bandit problem in which each distribution $G_i \in \mathcal{G}$ satisfies condition (2.1) of the Extremal Types Theorem. Consider the instance $\bar{I} = (\frac{n}{m}, \bar{\mathcal{G}})$, where $\bar{\mathcal{G}} = \{\bar{G}_1, \bar{G}_2, \ldots, \bar{G}_k\}$, and arm $\bar{G}_i$ returns the maximum payoff obtained by pulling the corresponding arm G_i m times. Effectively, $\bar{I}$ is a restricted version of I in which the arms must be pulled in batches of size m, rather than in any arbitrary order. For m sufficiently large, the Extremal Types Theorem guarantees that for each i, $\bar{G}_i \approxeq GEV_{(\mu_i,\sigma_i,\xi_i)}$ for some constants μ_i, σ_i, and ξ_i. Thus, the instance $I' = (\frac{n}{m}, \mathcal{G}')$ with $\mathcal{G}' = \{G'_1, G'_2, \ldots, G'_k\}$ and $G'_i = GEV_{(\mu_i,\sigma_i,\xi_i)}$ is approximately equivalent to $\bar{I}$ and satisfies our distributional assumptions.

The purpose of the restrictions on the parameters μ_i, σ_i, and ξ_i is to ensure that each GEV distribution has finite, bounded mean and variance.

4 An Asymptotically Optimal Algorithm

We will analyze the following max k-armed bandit strategy.

Strategy $\mathcal{S}^1(\epsilon, \delta)$:

1. *(Exploration)* For each arm $G_i \in \mathcal{G}$:

   (a) Using $t = O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ samples of G_i, obtain an estimate $\bar{m}_n^{G_i}$ of $m_n^{G_i}$. Assuming that arm G_i has shape parameter $\xi_i \leq 0$, our estimate will have the property that

   $$\mathbb{P}\left[|\bar{m}_n^{G_i} - m_n^{G_i}| < \epsilon\right] \geq 1 - \delta \,.$$

2. *(Exploitation)* Set $\hat{i} = \arg\max_{1 \leq i \leq k} \bar{m}_n^{G_i}$, and pull arm $G_{\hat{i}}$ for the remaining $n - tk$ trials.

If an arm G_i has shape parameter $\xi_i > 0$, the estimate obtained in step 1 (a) will instead have the property that $\mathbb{P}\left[\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < 1 + \epsilon\right] \geq 1 - \delta$ for constant α_1 independent of n.

The following theorem shows that with appropriate settings of ϵ and δ, strategy $\mathcal{S}^1$ is asymptotically optimal.

Theorem 1. *Let $I = (n, \mathcal{G})$ be an instance of the max k-armed bandit problem, where $\mathcal{G} = \{G_1, G_2, \ldots, G_k\}$ and $G_i = GEV_{(\mu_i, \sigma_i, \xi_i)}$. Let*

- $m_n^* = \max_{1 \leq i \leq k} m_n^{G_i}$,
- $\xi_{max} = \max_{1 \leq i \leq k} \xi_i$, *and*
- $S = \mathcal{S}^1\left(\sqrt[3]{\frac{k}{n}}, \frac{1}{kn^2}\right)$.

Then if $\xi_{max} \leq 0$,

$$\lim_{n \to \infty} S_n(I) = m_n^*$$

while if $\xi_{max} > 0$,

$$\lim_{n \to \infty} \frac{S_n(I)}{m_n^*} = 1 \,.$$

Proof.

Case $\xi_{max} \leq 0$. Let $\hat{m}_n = m_n^{G_{\hat{i}}}$ (where $\hat{i}$ is the arm selected for exploitation in step 2). Then $\hat{m}_{n-tk}$ is the expected maximum payoff obtained during the exploitation step, so

$$S_n(I) \geq \hat{m}_{n-tk} \,.$$

Claim 1. $\hat{m}_n - \hat{m}_{n-tk}$ is $O(\frac{tk}{n})$.

Proof of Claim 1. Let $\mu = \mu_{\hat{i}}$, $\sigma = \sigma_{\hat{i}}$, and $\xi = \xi_{\hat{i}}$ be the parameters of the arm selected for exploitation. Suppose $\xi = 0$. Then by Proposition 3, $\hat{m}_n - \hat{m}_{n-tk} = \sigma\left(\ln(n) - \ln(n - tk)\right)$.

Expanding $\ln(x+\beta)$ in powers of β about $\beta = 0$ for $|\beta| < \frac{x}{2}$, $x > 0$ yields

$$\begin{aligned} |\ln(x+\beta) - \ln(x)| &= \left| \sum_{i=1}^{\infty} (-1)^{i+1} \frac{1}{i} \left(\frac{\beta}{x}\right)^i \right| \\ &\leq \sum_{i=1}^{\infty} \left(\frac{|\beta|}{x}\right)^i \\ &= \frac{|\beta| x^{-1}}{1 - |\beta| x^{-1}} \\ &< 2\frac{|\beta|}{x} \end{aligned}$$

so for n sufficiently large, $\hat{m}_n - \hat{m}_{n-tk} \leq 2\sigma \frac{tk}{n} = O(\frac{tk}{n})$.

Now suppose $\xi < 0$. By Proposition 3, $\hat{m}_n - \hat{m}_{n-tk} = \frac{\sigma}{\xi}\Gamma(1-\xi)(n^\xi - (n-t)^\xi) = O((n-t)^\xi - n^\xi)$ where we have used the fact that $\frac{\sigma}{\xi}\Gamma(1-\xi) < 0$. Expanding $(n-t)^\xi$ in powers of t about $t = 0$ gives

$$(n-t)^\xi = n^\xi - \xi n^{\xi-1} t - O\left(\frac{t^2}{n^{\xi-2}}\right)$$

and so $(n-t)^\xi - n^\xi \leq -\xi n^{\xi-1} t \leq |\xi_\ell| \frac{t}{n} = O(\frac{t}{n})$. □

With probability at least $1 - k\delta$, all estimates obtained during the exploration phase are within ϵ of the correct values, so that $m_n^* - \hat{m}_n < 2\epsilon$. Assuming $m_n^* - \hat{m}_n < 2\epsilon$, it follows that

$$\begin{aligned} m_n^* - \hat{m}_{n-tk} &= (m_n^* - \hat{m}_n) + (\hat{m}_n - \hat{m}_{n-tk}) \\ &< 2\epsilon + O(\tfrac{tk}{n}) \\ &= 2\epsilon + \tfrac{k}{n} O\left(\ln(\tfrac{1}{\delta}) \tfrac{(\ln n)^2}{\epsilon^2}\right) \\ &= O(\Delta) \end{aligned}$$

where $\Delta = \ln(nk)\ln(n)^2 \sqrt[3]{\frac{k}{n}}$, and on the second line we have used Claim 1. Thus,

$$\begin{aligned} S_n(I) &\geq (1 - k\delta)(m_n^* - O(\Delta)) \\ &= m_n^* - O(\Delta) \\ &= m_n^* - o(1)\,. \end{aligned}$$

Case $\xi_{max} > 0$. See Appendix A.

□

Theorem 1 completes our analysis of the performance of $\mathcal{S}^1$. It remains only to describe how the estimates in step 1 (a) are obtained.

4.1 Estimating m_n

We adopt the following notation:

- Let $G = GEV_{\mu,\sigma,\xi}$ denote a GEV distribution with (unknown) parameters μ, σ, and ξ satisfying the conditions stated in §3, and

- let $m_i = m_i^G$.

To estimate m_n, we first obtain an accurate estimate of ξ. Then

1. if $\xi \approxeq 0$ (so that the growth of m_n as a function of $\ln n$ is linear), we estimate m_n by first estimating m_1 and m_2, then performing linear interpolation;
2. otherwise we estimate m_n by first estimating m_1, m_2, and m_4, then performing a nonlinear interpolation.

4.1.1 Estimating m_i for $i \in \{1, 2, 4\}$

The following two lemmas use well-known ideas to efficiently estimate m_i for small values of i.

Lemma 1. *Let i be a positive integer and let $\epsilon > 0$ be a real number. Then $O\left(\frac{i}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_i$ of m_i such that*

$$\mathbb{P}[|\bar{m}_i - m_i| < \epsilon] \geq \frac{3}{4}\,.$$

Proof. First consider the special case $i = 1$. Let X denote the sum of t draws from G, for some to-be-specified positive integer t. Then $\mathbb{E}[X] = m_1 t$ and $Var[X] = \tilde{\sigma}^2 t$, where $\tilde{\sigma}$ is the (unknown) standard deviation of G. We take $\bar{m}_1 = \frac{X}{t}$ as our estimate of m_1. Then

$$\begin{aligned}\mathbb{P}[|\bar{m}_1 - m_1| \geq \epsilon] &= \mathbb{P}[|t\bar{m}_1 - tm_1| \geq t\epsilon] \\ &= \mathbb{P}[|X - \mathbb{E}[X]| \geq \tfrac{\sqrt{t}\epsilon}{\tilde{\sigma}}\sqrt{Var[X]}] \\ &\leq \tfrac{\tilde{\sigma}^2}{t\epsilon^2}\end{aligned}$$

where the last inequality is Chebyshev's. Thus to guarantee $\mathbb{P}\left[|\bar{m}_1 - m_1| \geq \epsilon\right] \leq \frac{1}{4}$ we must set $t = \frac{4\tilde{\sigma}^2}{\epsilon^2} = O\left(\frac{1}{\epsilon^2}\right)$ (note that due to the assumptions in §3, $\tilde{\sigma}$ is $O(1)$).

For $i > 1$, we let X be the sum of t block maxima (each the maximum of i independent draws from G), which increases the number of samples required by a factor of i. □

To boost the probability that $|\bar{m}_i - m_i| < \epsilon$ from $\frac{3}{4}$ to $1 - \delta$, we use the "median of means" method.

Lemma 2. *Let i be a positive integer and let $\epsilon > 0$ and $\delta \in (0, 1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{i}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_i$ of m_i such that*

$$\mathbb{P}\left[|\bar{m}_n - m_n| < \epsilon\right] \geq 1 - \delta\,.$$

Proof. We invoke Lemma 1 r times (for r to be determined), yielding a set $E = \{\bar{m}_i^{(1)}, \bar{m}_i^{(2)}, \ldots, \bar{m}_i^{(r)}\}$ of estimates of m_i. Let $\bar{m}_i$ be the median element of E. Let $\mathcal{A} = \{\bar{m}_i^{(j)} \in E : |\bar{m}_i^{(j)} - m_i| < \epsilon\}$ be the set of "accurate" estimates of m_i; and let $A = |\mathcal{A}|$. Then $|\bar{m}_i - m_i| \geq \epsilon$ implies $A \leq \frac{r}{2}$, while $\mathbb{E}[A] \geq \frac{3}{4}r$. Using the standard Chernoff bound, we have

$$\mathbb{P}\left[|\bar{m}_i - m_i| \geq \epsilon\right] \leq \mathbb{P}\left[A \leq \frac{r}{2}\right] \leq \exp\left(-\frac{r}{C}\right)$$

for constant $C > 0$. Thus $r = O\left(\ln(\frac{1}{\delta})\right)$ repetitions suffice to ensure $\mathbb{P}[|\bar{m}_i - m_i| > \epsilon] \leq \delta$. □

4.1.2 Estimating m_n when $\xi = 0$

Lemma 3. *Assume G has shape parameter $\xi = 0$. Let n be a positive integer and let $\epsilon > 0$ and $\delta \in (0,1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_n$ of m_n such that*

$$\mathbb{P}[|\bar{m}_n - m_n| < \epsilon] \geq 1 - \delta \,.$$

Proof. By Proposition 3, $m_i = \mu + \sigma\gamma + \sigma\ln(i)$. Thus

$$m_n = m_1 + (m_2 - m_1)\log_2(n) \,. \tag{4.1}$$

Let $\bar{m}_1$ and $\bar{m}_2$ be estimates of m_1 and m_2, respectively, and let $\bar{m}_n$ be the estimate of m_n obtained by plugging $\bar{m}_1$ and $\bar{m}_2$ into (4.1). Define $\Delta_i = |\bar{m}_i - m_i|$ for $i \in \{1, 2, n\}$. Then

$$\Delta_n \leq (1 + \log_2(n))(\Delta_1 + \Delta_2) \,.$$

Thus to guarantee $\mathbb{P}[\Delta_n < \epsilon] \geq 1 - \delta$, it suffices that $\mathbb{P}\left[\Delta_i \leq \frac{\epsilon}{2(1+\log_2(n))}\right] \geq 1 - \frac{\delta}{2}$ for all $i \in \{1, 2\}$. By Lemma 2, this requires $O\left(\ln(\frac{1}{\delta})\frac{(\ln n)^2}{\epsilon^2}\right)$ draws from G. □

4.1.3 Estimating m_n when $\xi \neq 0$

For the purpose of the proofs presented here, we will make a minor assumption concerning an arm's shape parameter ξ_i: we assume that for some known constant $\xi^* > 0$,

$$|\xi_i| < \xi^* \Rightarrow \xi_i = 0 \,.$$

Removing this assumption does not fundamentally change the results, but it makes the proofs more complicated.

Lemma 5 shows how to efficiently estimate ξ. Lemmas 6 and 7 show how to efficiently estimate m_n in the cases $\xi < 0$ and $\xi > 0$, respectively. We will use the following lemma.

Lemma 4.

$$m_4 - m_2 \geq \tfrac{1}{4}\sigma \text{ and}$$
$$m_2 - m_1 \geq \tfrac{1}{8}\sigma \,.$$

Proof. See Appendix A. □

Lemma 5. *For real numbers $\epsilon > 0$ and $\delta \in (0,1)$, $O\left(\ln(\frac{1}{\delta})\frac{1}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{\xi}$ of ξ such that*

$$\mathbb{P}\left[|\bar{\xi} - \xi| < \epsilon\right] \geq 1 - \delta \,.$$

Proof. Using Proposition 3, it is straightforward to check that for any $\xi < 1$,

$$\xi = \log_2\left(\frac{m_4 - m_2}{m_2 - m_1}\right) \,. \tag{4.2}$$

Let $\bar{m}_1$, $\bar{m}_2$, and $\bar{m}_4$ be estimates of m_1, m_2, and m_4, respectively, and let $\bar{\xi}$ be the estimate of ξ obtained by plugging $\bar{m}_1$, $\bar{m}_2$, and $\bar{m}_4$ into (4.2). Define $\Delta_m = \max_{i\in\{1,2,4\}} |\bar{m}_i - m_i|$ and define $\Delta_\xi = |\bar{\xi} - \xi|$. We wish to upper bound Δ_ξ as a function of Δ_m.

In Claim 1 of Theorem 1 we showed that $|\ln(x+\beta) - \ln(x)| \leq 2\frac{\beta}{x}$ for $\beta \leq \frac{x}{2}$. Letting $N = m_4 - m_2$ and $D = m_2 - m_1$, and noting that $\xi = \log_2(N) - \log_2(D) = \frac{1}{\ln 2}(\ln(N) - \ln(D))$, it follows that

$$\Delta_\xi \leq \frac{1}{\ln(2)}\left(\frac{2(2\Delta_m)}{N} + \frac{2(2\Delta_m)}{D}\right)$$

for $\Delta_m < \frac{1}{2}\min(N, D)$. Thus by Lemma 4 and the assumption that $\sigma \geq \sigma_\ell$, Δ_ξ is $O(\Delta_m)$.

Define $\Delta_i = |\bar{m}_i - m_i|$, so that $\Delta_m = \max_{i\in\{1,2,4\}} \Delta_i$. Then to guarantee $\mathbb{P}[\Delta_\xi < \epsilon] \geq 1 - \delta$, it suffices that $\mathbb{P}\left[\Delta_i \leq \Omega(\epsilon)\right] \geq 1 - \frac{\delta}{3}$ for all $i \in \{1, 2, 4\}$. By Lemma 2, this requires $O\left(\ln(\frac{1}{\delta})\frac{1}{\epsilon^2}\right)$ draws from G.

□

Lemma 6. *Assume G has shape parameter $\xi \leq -\xi^*$. Let n be a positive integer and let $\epsilon > 0$ and $\delta \in (0, 1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{1}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_n$ of m_n such that*

$$\mathbb{P}[|\bar{m}_n - m_n| < \epsilon] \geq 1 - \delta\,.$$

Proof. By Proposition 3,

$$m_i = \mu + \frac{\sigma}{\xi}\left(i^\xi \Gamma(1-\xi) - 1\right)\,.$$

Define

$$\begin{aligned}\alpha_1 &= \mu - \sigma\xi^{-1}\\ \alpha_2 &= \sigma\xi^{-1}\Gamma(1-\xi)\\ \alpha_3 &= 2^\xi\end{aligned}$$

so that

$$m_i = \alpha_1 + \alpha_2\alpha_3^{\log_2(i)}\,. \tag{4.3}$$

Plugging in the values $i = 1$, $i = 2$, and $i = 4$ into (4.3) yields a system of three quadratic equations. Solving this system for α_1, α_2, and α_3 yields

$$\begin{aligned}\alpha_1 &= (m_1 m_4 - m_2^2)(m_1 - 2m_2 + m_4)^{-1}\\ \alpha_2 &= (-2m_1 m_2 + m_1^2 + m_2^2)(m_1 - 2m_2 + m_4)^{-1}\\ \alpha_3 &= (m_4 - m_2)(m_2 - m_1)^{-1}\,.\end{aligned}$$

Let $\bar{m}_1$, $\bar{m}_2$, and $\bar{m}_4$ be estimates of m_1, m_2, and m_4, respectively. Plugging $\bar{m}_1$, $\bar{m}_2$, and $\bar{m}_4$ into the above equations yields estimates, say $\bar{\alpha}_1$, $\bar{\alpha}_2$, and $\bar{\alpha}_3$, of α_1, α_2, and α_3, respectively. Define $\Delta_m = \max_{i\in\{1,2,4\}} |\bar{m}_i - m_i|$ and $\Delta_\alpha = \max_{i\in\{1,2,3\}} |\bar{\alpha}_i - \alpha_i|$. To complete the proof, we show that $|\bar{m}_n - m_n|$ is $O(\Delta_m)$. The argument consists of two parts: in claims 1 through 3 we show that Δ_α is $O(\Delta_m)$, then in Claim 4 we show that $|\bar{m}_n - m_n|$ is $O(\Delta_\alpha)$.

Claim 1. Each of the numerators in the expressions for α_1, α_2, and α_3 has absolute value bounded from above, while each of the denominators has absolute value bounded from below. (The bounds are independent of the unknown parameters of G.)

Proof of claim 1. The numerators will have bounded absolute value as long as m_1, m_2, and m_3 are bounded. Upper bounds on m_1, m_2, and m_3 follow from the restrictions on the parameters μ, σ, and ξ. As for the denominators, by Lemma 4 we have

$$\begin{aligned} |m_1 - 2m_2 + m_4| &= |(m_2 - m_1)(\alpha_3 - 1)| \\ &\geq \tfrac{1}{8}\sigma_\ell |2^{-\xi^*} - 1| \, . \end{aligned}$$

□

Claim 2. Let N and D be fixed real numbers, and let β_N and β_D be real numbers with $|\beta_D| < \frac{|D|}{2}$. Then $|\frac{N+\beta_N}{D+\beta_D} - \frac{N}{D}|$ is $O(|\beta_N| + |\beta_D|)$.

Proof of claim 2. First, using the Taylor series expansion of $\frac{N}{D+\beta_D}$,

$$\begin{aligned} |\tfrac{N}{D+\beta_D} - \tfrac{N}{D}| &= |\tfrac{N\beta_D}{D^2} \textstyle\sum_{i=0}^{\infty} (-1)^{i+1} \left(\frac{\beta_D}{D}\right)^i | \\ &\leq |\tfrac{N\beta_D}{D^2(1-\beta_D D^{-1})}| \\ &= O\left(|\beta_D|\right) \, . \end{aligned}$$

Then

$$\begin{aligned} \left|\tfrac{N+\beta_N}{D+\beta_D} - \tfrac{N}{D}\right| &\leq \left|\tfrac{N}{D+\beta_D} - \tfrac{N}{D}\right| + \left|\tfrac{\beta_N}{D+\beta_D}\right| \\ &= O\left(|\beta_N| + |\beta_D|\right) \, . \end{aligned}$$

□

Claim 3. Δ_α is $O(\Delta_m)$.

Proof of claim 3. We show that $|\bar{\alpha}_1 - \alpha_1|$ is $O(\Delta_m)$. Similar arguments show that $|\bar{\alpha}_2 - \alpha_2|$ and $|\bar{\alpha}_4 - \alpha_4|$ are $O(\Delta_m)$, which proves the claim. To see that $|\bar{\alpha}_1 - \alpha_1|$ is $O(\Delta_m)$, let $N = m_1 m_4 - m_2^2$, and let $D = m_1 - 2m_2 + m_4$, so that $\alpha_1 = \frac{N}{D}$. Define $\bar{N}$ and $\bar{D}$ in the natural way so that $\bar{\alpha}_1 = \frac{\bar{N}}{\bar{D}}$. Because m_1,m_2, and m_3 are all $O(1)$ (by Claim 1), it follows that both $|\bar{N} - N|$ and $|\bar{D} - D|$ are $O(\Delta_m)$. That $|\bar{\alpha}_1 - \alpha_1|$ is $O(\Delta_m)$ follows by Claim 2. □

Claim 4. $|\bar{m}_n - m_n|$ is $O\left(\Delta_\alpha\right)$.

Proof of claim 4. Because $\xi_\ell \leq \xi \leq -\xi^*$ it must be that $0 < 2^{\xi_\ell} < \alpha_3 < 2^{-\xi^*} < 1$. So for Δ_α sufficiently small, $0 < \bar{\alpha}_3 < 1$.

$$\begin{aligned} |\bar{m}_n - m_n| &= |(\bar{\alpha}_1 + \bar{\alpha}_2 \bar{\alpha}_3^{\log_2(n)}) - (\alpha_1 + \alpha_2 \alpha_3^{\log_2(n)})| \\ &\leq |\bar{\alpha}_1 - \alpha_1| + |\bar{\alpha}_2 \bar{\alpha}_3^{\log_2(n)} - \bar{\alpha}_2 \alpha_3^{\log_2(n)}| \\ &\qquad + |\bar{\alpha}_2 \alpha_3^{\log_2(n)} - \alpha_2 \alpha_3^{\log_2(n)}| \\ &\leq |\bar{\alpha}_1 - \alpha_1| + |\bar{\alpha}_2||\bar{\alpha}_3 - \alpha_3| + |\bar{\alpha}_2 - \alpha_2| \\ &= O(\Delta_\alpha) \end{aligned}$$

where on the third line we have used the fact that both α_3 and $\bar{\alpha}_3$ are between 0 and 1, and in the last line we have used the fact that $|\bar{\alpha}_2|$ is $O(1)$. □

Putting claims 3 and 4 together, $|\bar{m}_n - m_n|$ is $O(\Delta_m)$. Define $\Delta_i = |\bar{m}_i - m_i|$, so that $\Delta_m = \max_{i \in \{1,2,4\}} \Delta_i$. Thus to guarantee $\mathbb{P}[|\bar{m}_n - m_n| < \epsilon] \geq 1 - \delta$, it suffices that $\mathbb{P}\left[\Delta_i < \Omega(\epsilon)\right] \geq 1 - \frac{\delta}{3}$ for all $i \in \{1, 2, 4\}$. By Lemma 2, this requires $O(\ln(\frac{1}{\delta})\frac{1}{\epsilon^2})$ draws from G. □

Lemma 7. *Assume G has shape parameter $\xi \geq \xi^*$. Let n be a positive integer and let $\epsilon > 0$ and $\delta \in (0, 1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_n$ of m_n such that*

$$\mathbb{P}\left[\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < (1+\epsilon)\right] \geq 1 - \delta$$

where $\alpha_1 = \mu - \frac{\sigma}{\xi}$.

Proof. See Appendix A. □

Putting the results of lemmas 3, 5, 6, and 7 together, we obtain the following theorem.

Theorem 2. *Let n be a positive integer and let $\epsilon > 0$ and $\delta \in (0, 1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_n$ of m_n such that with probability at least $1 - \delta$, one of the following holds:*

- *$\xi \leq 0$ and $|\bar{m}_n - m_n| < \epsilon$, or*
- *$\xi > 0$ and $\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < 1 + \epsilon$, where $\alpha_1 = \mu - \frac{\sigma}{\xi}$.*

Proof. First, invoke Lemma 5 with parameters $\frac{\xi^*}{3}$ and $\frac{\delta}{2}$. Then invoke one of Lemmas 3, 6, or 7 (depending on the estimate $\bar{\xi}$ obtained from Lemma 5) with parameters ϵ and $\frac{\delta}{2}$. □

Theorem 2 shows that step 1 (a) of strategy $\mathcal{S}^1$ can be performed as described.

5 Conclusions

The max k-armed bandit problem is a variant of the classical k-armed bandit problem with practical applications to combinatorial optimization. Motivated by extreme value theory, we studied a restricted version of this problem in which each arm yields payoff drawn from a GEV distribution. We derived PAC bounds on the sample complexity of estimating m_n, the expected maximum of n draws from a GEV distribution. Using these bounds, we showed that a simple algorithm for the max k-armed bandit problem is asymptotically optimal. Ours is the first algorithm for this problem with rigorous asymptotic performance guarantees.

References

[1] Donald. A. Berry and Bert Fristedt. *Bandit Problems: Sequential Allocation of Experiments*. Chapman and Hall, London, 1986.

[2] Vincent A. Cicirello and Stephen F. Smith. Heuristic selection for stochastic search optimization: Modeling solution quality by extreme value theory. In *Proceedings of the 10th International Conference on Principles and Practice of Constraint Programming*, pages 197–211, 2004.

[3] Vincent A. Cicirello and Stephen F. Smith. The max k-armed bandit: A new model of exploration applied to search heuristic selection. In *Proceedings of AAAI 2005*, pages 1355–1361, 2005.

[4] Stuart Coles. *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London, 2001.

[5] Philip W. L. Fong. A quantitative study of hypothesis selection. In *International Conference on Machine Learning*, pages 226–234, 1995.

[6] Leslie P. Kaelbling. *Learning in Embedded Systems*. The MIT Press, Cambridge, MA, 1993.

[7] Herbert Robbins. Some aspects of sequential design of experiments. *Bulletin of the American Mathematical Society*, 58:527–535, 1952.

Appendix A

Theorem 1. *Let $I = (n, \mathcal{G})$ be an instance of the max k-armed bandit problem, where $\mathcal{G} = \{G_1, G_2, \ldots, G_k\}$ and $G_i = GEV_{(\mu_i,\sigma_i,\xi_i)}$. Let*

- $m_n^* = \max_{1\leq i\leq k} m_n^{G_i}$,
- $\xi_{max} = \max_{1\leq i\leq k} \xi_i$, *and*
- $S = \mathcal{S}^1\left(\sqrt[3]{\frac{k}{n}}, \frac{1}{kn^2}\right)$.

Then if $\xi_{max} \leq 0$,

$$\lim_{n\to\infty} S_n(I) = m_n^*$$

while if $\xi_{max} > 0$,

$$\lim_{n\to\infty} \frac{S_n(I)}{m_n^*} = 1\,.$$

Proof. For $\xi_{max} \leq 0$, the theorem was proved in the main text. It remains only to address the case $\xi_{max} > 0$.

Case $\xi_{max} > 0$. We will prove the stronger claim that

$$\frac{S_n(I) - \alpha_1}{m_n^* - \alpha_1} = 1 - O(\Delta) \tag{5.1}$$

where $\Delta = \ln(nk)\ln(n)^2\sqrt[3]{\frac{k}{n}}$ and $\alpha_1 = \max_{1\leq i\leq k} \alpha_i$, where $\alpha_i = \mu_i - \frac{\sigma_i}{\xi_i}$.

For the moment, let us assume that *all* arms have shape parameter $\xi_i > 0$. Let $\mathcal{A}$ be the event (which occurs with probability at least $1 - k\delta$) that all estimates obtained in step 1 (a) satisfy the inequality in Theorem 2.

Claim 1. To prove (5.1), it suffices to show that $\mathcal{A}$ implies $\frac{m_n^* - \alpha_1}{\hat{m}_{n-tk} - \alpha_1} = 1 + O(\Delta)$.

Proof of claim 1. Because $S_n(I) \geq \hat{m}_{n-tk}$ and the event $\mathcal{A}$ occurs with probability at least $1 - k\delta$, it suffices to show that $\mathcal{A}$ implies

$$\frac{(1 - \delta k)\hat{m}_{n-tk} - \alpha_1}{m_n^* - \alpha_1} = 1 - O(\Delta) \, .$$

Because $\frac{\delta k(\hat{m}_{n-tk})}{m_n^* - \alpha_1}$ is $O\left(\frac{1}{n^2}\right) = o(\Delta)$, it suffices to show that $\mathcal{A}$ implies

$$\frac{\hat{m}_{n-tk} - \alpha_1}{m_n^* - \alpha_1} = 1 - O(\Delta) \, .$$

This can be rewritten as $m_n^* - \alpha_1 = (\hat{m}_{n-tk} - \alpha_1)\frac{1}{1-O(\Delta)} = (\hat{m}_{n-tk} - \alpha_1)(1 + O(\Delta))$ (we can replace $\frac{1}{1-O(\Delta)}$ with $1 + O(\Delta)$ because for $r < \frac{1}{2}$, $\frac{1}{1-r} = 1 + \frac{r}{1-r} < 1 + 2r$). □

Claim 2. $\frac{\hat{m}_n - \hat{\alpha}_1}{\hat{m}_{n-tk} - \hat{\alpha}_1} = 1 + O(\Delta)$.

Proof of claim 2. Using Proposition 3,

$$\begin{aligned} \ln\left(\tfrac{\hat{m}_n - \hat{\alpha}_1}{\hat{m}_{n-tk} - \hat{\alpha}_1}\right) &= \ln\left(\tfrac{n^\xi}{(n-t)^\xi}\right) \\ &= \xi\left(\ln(n) - \ln(n - tk)\right) \\ &= O(\tfrac{tk}{n}) \\ &= O(\Delta) \end{aligned}$$

The claim follows from the fact that $\exp(\beta) < 1 + \frac{3}{2}\beta$ for $\beta < \frac{1}{2}$, so that $\exp(O(\Delta)) = 1 + O(\Delta)$. □

Claim 3. $\mathcal{A}$ implies that for all i,

$$\frac{\bar{m}_n^i - \alpha_1}{m_n^i - \alpha_1} < 1 + \epsilon \, .$$

Proof of claim 3. By definition, $\alpha_1 = \alpha_1^i - \beta$ for some $\beta \geq 0$. The claim follows from the fact that for positive N and D and $\beta \geq 0$, $\frac{N}{D} < 1 + \epsilon$ implies $\frac{N+\beta}{D+\beta} < 1 + \epsilon$. □

Claim 4. $\mathcal{A}$ implies $\frac{m_n^* - \alpha_1}{\hat{m}_{n-tk} - \alpha_1} = 1 + O(\Delta)$.

Proof of claim 4.

$$\begin{aligned} \tfrac{m_n^* - \alpha_1}{m_{\hat{n}-tk}^{\hat{i}} - \alpha_1} &= \tfrac{m_n^* - \alpha_1}{\bar{m}_n^* - \alpha_1} \cdot \tfrac{\bar{m}_n^* - \alpha_1}{\bar{m}_n^{\hat{i}} - \alpha_1} \cdot \tfrac{\bar{m}_n^{\hat{i}} - \alpha_1}{m_n^{\hat{i}} - \alpha_1} \cdot \tfrac{m_n^{\hat{i}} - \alpha_1}{m_{\hat{n}-tk}^{\hat{i}} - \alpha_1} \\ &\leq (1 + \epsilon) \cdot 1 \cdot (1 + \epsilon) \cdot (1 + O(\Delta)) \\ &= 1 + O(\Delta) \end{aligned}$$

where in the second step we have used claims 2 and 3. □

Putting claims 1 and 4 together completes the proof. To remove the assumption that all arms have $\xi_i > 0$, we need to show that $\mathcal{A}$ implies that for n sufficiently large, the arms $\hat{i}$ and i^* (the only arms that play a role in the proof) will have shape parameters > 0. This follows from the fact that if $\xi_i \leq 0$, m_n^i is $O(\ln(n))$, while if $\xi_i > \xi^* > 0$, m_n^i is $\Omega(n^{\xi_i})$.

□

Lemma 4.

$$m_4 - m_2 \geq \tfrac{1}{4}\sigma \text{ and}$$
$$m_2 - m_1 \geq \tfrac{1}{8}\sigma \,.$$

Proof. If $\xi = 0$, then by Proposition 3, $m_4 - m_2 = m_2 - m_1 = \ln(2)\sigma$ and we are done. Otherwise,

$$m_4 - m_2 = \sigma(2^\xi - 1)\xi^{-1}\Gamma(1-\xi) \text{ and}$$
$$m_2 - m_1 = \sigma(4^\xi - 2^\xi)\xi^{-1}\Gamma(1-\xi)\,.$$

It thus suffices to prove the following claim.

Claim 1.

$$\min_{\xi < \frac{1}{2}} \left\{ \frac{2^\xi - 1}{\xi}\Gamma(1-\xi) \right\} \geq \frac{1}{4}, \text{ and}$$
$$\min_{\xi < \frac{1}{2}} \left\{ \frac{4^\xi - 2^\xi}{\xi}\Gamma(1-\xi) \right\} \geq \frac{1}{8}\,.$$

Proof of claim 1. We state without proof the following properties of the Γ function:

$$\Gamma(z) \geq \lfloor z \rfloor! \quad \forall z \geq 2$$
$$\Gamma(z) \geq \tfrac{1}{2} \quad \forall z > 0$$

Making the change of variable $y = -\xi$, it suffices to show

$$\min_{y > -\frac{1}{2}} \left\{ \frac{1 - 2^{-y}}{y}\Gamma(1+y) \right\} \geq \frac{1}{4}\,, \text{ and} \tag{5.2}$$

$$\min_{y > -\frac{1}{2}} \left\{ \frac{2^{-y}(1 - 2^{-y})}{y}\Gamma(1+y) \right\} \geq \frac{1}{8}\,. \tag{5.3}$$

(5.2) holds because for $-\frac{1}{2} < y \leq 1$,

$$\frac{1 - 2^{-y}}{y}\Gamma(1+y) \geq \frac{1}{2}\Gamma(1+y) \geq \frac{1}{4}\,,$$

while for $y > 1$,

$$\frac{1 - 2^{-y}}{y}\Gamma(1+y) \geq \frac{\lfloor y+1 \rfloor!}{2y} \geq \frac{1}{2}\,.$$

Similarly, (5.3) holds because for $-\frac{1}{2} < y \leq 1$,

$$\frac{2^{-y}(1 - 2^{-y})}{y}\Gamma(1+y) \geq \frac{1}{8}\,,$$

while for $y > 1$,

$$\frac{2^{-y}(1-2^{-y})}{y}\Gamma(1+y) \geq \frac{\lfloor y+1 \rfloor !}{2y(2^y)} \geq \frac{1}{8}\,.$$

□

□

Lemma 7. *Assume G has shape parameter $\xi \geq \xi^*$. Let n be a positive integer and let $\epsilon > 0$ and $\delta \in (0,1)$ be real numbers. Then $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ draws from G suffice to obtain an estimate $\bar{m}_n$ of m_n such that*

$$\mathbb{P}\left[\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < (1+\epsilon)\right] \geq 1-\delta$$

where $\alpha_1 = \mu - \frac{\sigma}{\xi}$.

Proof. We use the same estimation procedure as in the proof of Lemma 6. Let α_1, α_2, α_3, Δ_α, and Δ_m be defined as they were in that proof.

The inequality $\frac{1}{1+\epsilon} < \frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1} < 1+\epsilon$ is the same as $|\ln(\frac{\bar{m}_n - \alpha_1}{m_n - \alpha_1})| < \ln(1+\epsilon)$. For $\epsilon < \frac{1}{2}$, $\ln(1+\epsilon) \geq \frac{7}{8}\epsilon$, so it suffices to guarantee that

$$|\ln(\bar{m}_n - \alpha_1) - \ln(m_n - \alpha_1)| < \frac{7}{8}\epsilon\,.$$

Claim 1. $|\ln(\bar{m}_n - \alpha_1) - \ln(m_n - \alpha_1)|$ is $O(\ln(n)\Delta_\alpha)$.

Proof of claim 1. Because $|\ln(\bar{m}_n - \alpha_1) - \ln(\bar{m}_n - \bar{\alpha}_1)|$ is $O(\Delta_\alpha)$, it suffices to show that $|\ln(\bar{m}_n - \bar{\alpha}_1) - \ln(m_n - \alpha_1)|$ is $O(\ln(n)\Delta_\alpha)$. This is true because

$$\begin{aligned}
\ln(\bar{m}_n - \bar{\alpha}_1) &= \ln\left(\bar{\alpha}_2 \bar{\alpha}_3^{\log_2(n)}\right) \\
&= \log_2(n)\ln(\bar{\alpha}_3) + \ln(\bar{\alpha}_2) \\
&= \log_2(n)\ln(\alpha_3) + \ln(\alpha_2) \pm O\left(\ln(n)\Delta_\alpha\right) \\
&= \ln\left(\alpha_2 \alpha_3^{\log_2(n)}\right) \pm O\left(\ln(n)\Delta_\alpha\right) \\
&= \ln(m_n - \alpha_1) \pm O\left(\ln(n)\Delta_\alpha\right)\,.
\end{aligned}$$

□

Setting $\Delta_\alpha < \Omega\left(\ln(n)^{-1}\epsilon\right)$ then guarantees $|\ln(\bar{m}_n) - \ln(m_n)| < \frac{7}{8}\epsilon$. By Claim 3 of the proof of Lemma 6 (which did not depend on the assumption $\xi < 0$), Δ_α is $O\left(\Delta_m\right)$, so we require $\mathbb{P}[\Delta_m < \Omega\left(\ln(n)^{-1}\epsilon\right)] \geq 1-\delta$. Define $\Delta_i = |\bar{m}_i - m_i|$, so that $\Delta_m = \max_{i\in\{1,2,4\}} \Delta_i$. It suffices that $\mathbb{P}\left[\Delta_i < \Omega\left(\ln(n)^{-1}\epsilon\right)\right] \geq 1 - \frac{\delta}{3}$ for $i \in \{1,2,4\}$. By Lemma 2, ensuring this requires $O\left(\ln(\frac{1}{\delta})\frac{\ln(n)^2}{\epsilon^2}\right)$ draws from G.

□