Quantifying Human Perception with Multi-Armed Bandits

Extended Abstract

Julien Audiffren
Cognition and Perception Laboratory, University of Fribourg
Fribourg, Switzerland
julien.audiffren@unifr.ch

ABSTRACT

We study a variant of the continuous multi-armed bandits problem, where the objective is to estimate the sensitivity threshold for an unknown psychometric function Ψ. This setting models the conduct of a psychometric experiment, which aims at quantifying human perception. We show that this setting is akin to hierarchical multi-armed bandits and Black-box optimization of noisy functions, with both significant similarities and key differences. We introduce a new algorithm, DOS, for Dichotomous Optimistic Search, that efficiently solves this task, and show that DOS outperforms recent methods from both Psychophysics and Global Optimization for non Gaussian Psychometric functions in our experiments.[1]

KEYWORDS

Multi-armed Bandits, Noisy Black Box Optimization, Psychophysics

ACM Reference Format:
Julien Audiffren . 2021. Quantifying Human Perception with Multi-Armed Bandits: Extended Abstract. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 3 pages.

1 INTRODUCTION

Psychophysics investigates the relation between physical stimuli and the subjective responses (such as sensations) they produce. One of the key aspect of Psychophysics is the evaluation of human perception, which is generally assessed by performing psychometric experiments, which unfold as follows: the experimenter presents to an individual, called the observer, a sequence of stimuli of varying intensities (for instance, the volume of a specific sound, see e.g. [14, 20]), and try to measure how often the different intensities are perceived by the observer. In particular, the majority of experiments are interested in measuring the *sensitivity threshold*, where the stimulus is just noticeable [28]. In the recent years, there has been an increased interest in using adaptive algorithms in order to estimate this threshold in psychometric experiments [30], where an agent adapts the sequence of stimulus intensity based on the observer responses. The two most popular adaptive methods in Psychophysics are currently the staircase [12] and likelihood maximization [28], both of which require strong assumption regarding the psychometric function and have limited guarantees regarding the consistency of their estimator.

In this paper, we show that the threshold estimation problem can be rewritten as a new type of the pure exploration continuous multi-armed bandit problem, with interesting twists. Then, we introduce a new algorithm, Dichotomous Optimistic Search (DOS), that takes inspiration from hierarchical bandits and black box optimization (see e.g. [18, 38]) to solve this problem (Section 3). The idea behind DOS is to perform a stochastic continuous binary search, while achieving the correct trade off between the depth of the binary tree, and the confidence in its noisy comparisons. DOS only assumes a minimal set of hypotheses over the psychometric function, and does not assume the knowledge of its shape. Our experiments show that DOS significantly outperforms traditional adaptive psychometric methods and recent global optimization methods.

2 PROBLEM SETUP

Let T denote the time horizon, $\mathbb{I} = [0, 1]$ the interval of possible stimuli [2], $\Psi : \mathbb{I} \mapsto [0, 1]$ the psychometric function, $\mu_* \in [0, 1]$ the target probability, $s_* \doteq \Psi^{-1}(\mu_*)$ the sensitivity threshold. Due to the nature of the task, the psychometric function is assumed to be continuous and strictly increasing (see e.g. [30]). The objective of the threshold estimation problem is to find an estimator $\hat{s}$ of the sensitivity threshold s_* with at most T stimuli. $\mathbb{I}$, T and μ_* are known to the agent (here the experimenter), but Ψ is unknown. The process unfolds as follows. For each round $t \in [1, \dots, T]$:

(1) The agent chooses an arm (here an intensity) $s \in \mathbb{I}$.
(2) The environment (here the observer) detects the stimulus and notify the agent using an independent Bernoulli random variable of mean $\Psi(s)$.

At time $t = T$, the agent returns the arm $\hat{s}$ that is her best guess for the target stimulus s_*. The performance of the agent is then evaluated using simple regret $\mathcal{R}$, defined as $\mathcal{R}(\hat{s}) = |\mu_* - \Psi(\hat{s})|$. In the rest of the paper, we make the following assumption on Ψ.

Assumption 1 (Ψ is smooth around s_*)**.** There exists $\nu > 0$, and $0 < \rho < 1$ such that $\forall h > 0$, $\forall s \in \mathbb{I}$, $\quad |s - s_*| \leq 2^{-h} \implies |\Psi(s) - \Psi(s_*)| \leq \nu\rho^h$

This implies that Ψ is smooth enough around s_* to prevent the "find the needle in a haystack" problem of global optimization [43]. It should be noted that all continuously differentiable Ψ (including e.g. Gaussian c.d.f.) satisfy Assumption 1.

Relation with Global Optimization. Let f be defined as $f(s) = -|\mu_* - \Psi(s)|$. It is easy to see that f admits s_* as its unique maximum, and $f(s_*) = 0$. Moreover, the regret defined above is equivalent to the usual definition of simple regret for f (see e.g. [7]). Similarly, Assumption 1 implies a similar smoothness condition for f around

[1]More elements of analysis of DOS and additional experiments may be found in [2].

[2]The present work can be easily extended to any closed interval

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.

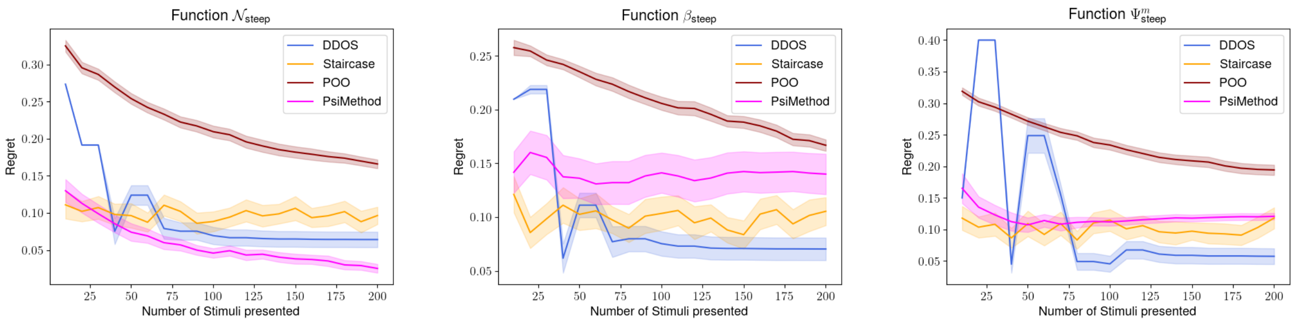

Figure 1: Comparison of the evolution of the average regret over 100 runs as a function of the number of stimuli presented to the observer, for a time horizon of $T = 200$, for each psychometric function. The standard deviation is reported using the shaded area.

Algorithm 1: DOS

Parameters μ_* (objective), T (time horizon)
Initialization $i \leftarrow 1$, $s_1 \leftarrow 1/2$, $N_1 \leftarrow 0$, $\hat{\mu}_1 \leftarrow 0$, $t \leftarrow 0$, $\mathcal{S} = \emptyset$, N_* as in (3) and $\mathcal{B}_T(\cdot)$ as in (2).
while $t \leq T$ **do**
  if $|\mu_* - \hat{\mu}_i(t)| > 2\mathcal{B}_T(N_i(t))$ ***or*** $N_i(t) > N_*$ **then**
    if $N_i(t) > N_*$ **then**
      $\mathcal{S} \leftarrow \mathcal{S} \cup \{i\}$;
    Activate new arm using (1) ;
    $i \leftarrow i + 1$;
  Sample arm s_i, update t, N_i, $\hat{\mu}_i$;
Output: s_{i_*}, where $i_* = \max \mathcal{S}$ if $\mathcal{S} \neq \emptyset$, else i.

its maximum. Therefore, f draws a link between black box optimization and threshold estimation. However, since Ψ is unknown and only observed through the realizations of Bernoulli random variables, global optimization strategies cannot be directly used to solve the threshold optimization problem.

3 DICHOTOMOUS OPTIMISTIC SEARCH

We now introduce our main contribution, DOS. Let s_i denotes the stimulus value of the i-th arm activated by DOS, and $N_i(t)$ (resp. $\hat{\mu}_i(t)$ and μ_i) the number of pulls (resp. the empirical average and the true probability value) of the i-th arm at time t.

DOS strategy. The pseudocode for DOS can be found in Algorithm 1. The general idea of DOS is inspired by the deterministic dichotomous search algorithm. In order to achieve this, the agent starts with the arm $s_1 = 1/2$ (i.e. the center of $\mathbb{I}$). Then, the agent pulls the latest arm of the sequence s_i until the time budget is elapsed ($t = T$) or one of the two possible new arm activation criteria is satisfied. Then she compares μ_*, the target probability, and $\hat{\mu}_i(N_i)$, (i.e. the empirical proportion of stimuli of intensity s_i that were detected). Finally, leveraging the fact that Ψ is monotonically increasing, she activates the arm s_{i+1} such that

$$s_{i+1} = s_i + \text{sign}(\mu_* - \hat{\mu}_i)(1/2^{i+1}) \tag{1}$$

Contrarily to the deterministic setting, here the agent has only access to noisy observations of $\Psi(s_i)$. Therefore, for any arm s_i the agent can only compare $\hat{\mu}_i$ and μ_*, and can never be sure if $\Psi(s_i) \geq \mu_*$. To succeed, the agent maintains a trade-off between:

- **Confidence**: Increase N_i to improve confidence in the $\hat{\mu}_i$,
- **Depth**: Increase i to improve the approximation of s_*.

This is achieved by using two arm activation rules:

$$|\mu_* - \hat{\mu}_i(t)| > \mathcal{B}_T(N_i(t)) \doteq \frac{3}{2}\sqrt{\frac{\log(T)}{N_i(t)}}. \tag{2}$$

$$N_i(t) > N_* \doteq \left\lfloor \frac{T}{(\log T)(\log^2 T)} \right\rfloor. \tag{3}$$

While (2) is a direct application of Azuma Hoeffding, (3) is key in achieving the aforementioned exploration trade-off, by limiting number of pulls for the arm i before the activation of the next arm.

4 EXPERIMENTS

We now evaluate the performance of DOS. We set $T = 200$, to reproduce the constraints of psychometric experiments. We compare DOS to the two commonly used adaptive methods in Psychophysics: Staircase [12], and PsiMethod [28], and to the hierarchical bandit based algorithm POO (Parallel Optimistic Optimization – [18]). The objective is to identify the stimulus s_* such that $\mu_* = 0.5$. We used three psychometric functions : $\mathcal{N}_{\text{steep}}$, based on a Gaussian c.d.f., β_{steep}, are based on a Beta c.d.f., and Ψ^m_{steep}, defined as :

$$\Psi^m(s^* + x) = \min\left(1, \mu_* + |x|\right) 1_{x \geq 0} + \max\left(0, \mu_* - |x|^{0.3}\right) 1_{x \leq 0}$$

Results. Figure 1 reports the average simple regret over 100 runs for each method and psychometric function. First, note that PsiMethod outperforms other algorithms for $\mathcal{N}_{\text{steep}}$ – as it is able to leverage its additional assumption about the Gaussian c.d.f. – and this advantage is particularly important for small time budget. However, PsiMethod performs poorly for the other psychometric functions, that are non Gaussian. Second, while POO seems to converge toward the solution for every function, it achieves the worst performance, as its rate of convergence is slow and it cannot take advantage of the monotonic property of Ψ. Finally, it is important to note that DOS provides one of the best estimation in all these settings.

REFERENCES

[1] Rocío Alcalá-Quintana and Miguel García-pérez. 2005. Stopping rules in Bayesian adaptive threshold estimation. *Spatial Vision* 18, 3 (2005), 347–374. https://doi.org/10.1163/1568568054089375

[2] Julien Audiffren. 2021. Multi-Armed Bandits for Quantifying Human Perception with Dichotomous Optimistic Search. (Jan. 2021). https://hal.archives-ouvertes.fr/hal-02448282 working paper or preprint.

[3] Peter Auer, Ronald Ortner, and Csaba Szepesvári. 2007. Improved Rates for the Stochastic Continuum-Armed Bandit Problem. In *Learning Theory (Lecture Notes in Computer Science)*, Nader H. Bshouty and Claudio Gentile (Eds.). Springer, Berlin, Heidelberg, 454–468. https://doi.org/10.1007/978-3-540-72927-3_33

[4] Maryam Aziz, Jesse Anderton, Emilie Kaufmann, and Javed Aslam. 2018. Pure Exploration in Infinitely-Armed Bandit Models with Fixed-Confidence. In *Algorithmic Learning Theory*. 3–24. http://proceedings.mlr.press/v83/aziz18a.html

[5] A. J. Benson, E. C. Hutt, and S. F. Brown. 1989. Thresholds for the perception of whole body angular movement about a vertical axis. *Aviation, Space, and Environmental Medicine* 60, 3 (1989), 205–213.

[6] Lawrence G Brown. 1996. Additional rules for the transformed up-down method in psychophysics. *Perception & Psychophysics* 58, 6 (1996), 959–962.

[7] Sébastien Bubeck, Rémi Munos, and Gilles Stoltz. 2011. Pure exploration in finitely-armed and continuous-armed bandits. *Theoretical Computer Science* 412, 19 (April 2011), 1832–1852. https://doi.org/10.1016/j.tcs.2010.12.059

[8] Séebastian Bubeck, Tengyao Wang, and Nitin Viswanathan. 2013. Multiple Identifications in Multi-Armed Bandits. In *International Conference on Machine Learning*. 258–265. http://proceedings.mlr.press/v28/bubeck13.html

[9] Adam D. Bull. 2015. Adaptive-treed bandits. *Bernoulli* 21, 4 (Nov. 2015), 2289–2307. https://doi.org/10.3150/14-BEJ644

[10] Arghya Roy Chaudhuri and Shivaram Kalyanakrishnan. 2017. PAC Identification of a Bandit Arm Relative to a Reward Quantile. In *Thirty-First AAAI Conference on Artificial Intelligence*. https://www.aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14335

[11] Emile Contal, David Buffoni, Alexandre Robicquet, and Nicolas Vayatis. 2013. Parallel Gaussian process optimization with upper confidence bound and pure exploration. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*. Springer, 225–240.

[12] Tom N. Cornsweet. 1962. The Staircase-Method in Psychophysics. *The American Journal of Psychology* 75, 3 (Sept. 1962), 485. https://doi.org/10.2307/1419876

[13] Florin Cutzu and John K Tsotsos. 2003. The selective tuning model of attention: psychophysical evidence for a suppressive annulus around an attended item. *Vision research* 43, 2 (2003), 205–219.

[14] WOUTER A Dreschler and J Verschuure. 1996. *Psychophysical evaluation of fast compression systems*. World Scientific Singapore.

[15] Xavier Fontaine, Shie Mannor, and Vianney Perchet. 2020. An adaptive stochastic optimization algorithm for resource allocation. In *Algorithmic Learning Theory*. 319–363.

[16] Ingo Fründ, N. Valentin Haenel, and Felix A. Wichmann. 2011. Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision* 11, 6 (May 2011), 16–16. https://doi.org/10.1167/11.6.16

[17] Miguel A. García-Pérez and Rocío Alcalá-Quintana. 2007. Bayesian adaptive estimation of arbitrary points on a psychometric function. *Brit. J. Math. Statist. Psych.* 60, 1 (2007), 147–174. https://doi.org/10.1348/000711006X104596

[18] Jean-Bastien Grill, Michal Valko, and Rémi Munos. 2015. Black-box optimization of noisy functions with unknown smoothness. https://hal.inria.fr/hal-01222915

[19] Christian Hatzfeld, Viet Quoc Hoang, and Mario Kupnik. 2016. It's All About the Subject - Options to Improve Psychometric Procedure Performance. In *Haptics: Perception, Devices, Control, and Applications*, Fernando Bello, Hiroyuki Kajimoto, and Yon Visell (Eds.). Vol. 9774. Springer International Publishing, Cham, 394–403. https://doi.org/10.1007/978-3-319-42321-0-36

[20] Tatsuya Hirahara. 2004. Physical characteristics of headphones used in psychophysical experiments. *Acoustical science and technology* 25, 4 (2004), 276–285.

[21] Shivaram Kalyanakrishnan, Ambuj Tewari, Peter Auer, and Peter Stone. 2012. PAC Subset Selection in Stochastic Multi-armed Bandits.

[22] Faisal Karmali, Shomesh E. Chaudhuri, Yongwoo Yi, and Daniel M. Merfeld. 2016. Determining thresholds using adaptive procedures and psychometric fits: evaluating efficiency using theory, simulations, and human experiments. *Experimental brain research* 234, 3 (March 2016), 773–789. https://doi.org/10.1007/s00221-015-4501-8

[23] Emilie Kaufmann and Aurélien Garivier. 2017. Learning the distribution with largest mean: two bandit frameworks. *ESAIM: Proceedings and Surveys* 60 (2017), 114–131. https://doi.org/10.1051/proc/201760114

[24] Emilie Kaufmann and Shivaram Kalyanakrishnan. 2013. Information Complexity in Bandit Subset Selection. In *Conference on Learning Theory*. 228–251. http://proceedings.mlr.press/v30/Kaufmann13.html

[25] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. 2008. Multi-armed bandits in metric spaces. In *Proceedings of the fortieth annual ACM symposium on Theory of computing*. 681–690.

[26] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. 2019. Bandits and experts in metric spaces. *Journal of the ACM (JACM)* 66, 4 (2019), 1–77.

[27] Robert Kleinberg, Aleksandrs Slivkins, and Eli Upfal. 2019. Bandits and Experts in Metric Spaces. *arXiv:1312.1277 [cs]* (April 2019). http://arxiv.org/abs/1312.1277 arXiv: 1312.1277.

[28] Leonid L. Kontsevich and Christopher W. Tyler. 1999. Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research* 39, 16 (Aug. 1999), 2729–2737. https://doi.org/10.1016/S0042-6989(98)00285-5

[29] Malte Kuss, Frank Jäkel, and Felix A. Wichmann. 2005. Bayesian inference for psychometric functions. *Journal of Vision* 5, 5 (May 2005), 8–8. https://doi.org/10.1167/5.5.8

[30] Marjorie R. Leek. 2001. Adaptive procedures in psychophysical research. *Perception & Psychophysics* 63, 8 (Nov. 2001), 1279–1292. https://doi.org/10.3758/BF03194543

[31] Gábor Lengyel and József Fiser. 2019. The relationship between initial threshold, learning, and generalization in perceptual learning. *Journal of Vision* 19, 4 (April 2019), 28–28. https://doi.org/10.1167/19.4.28

[32] H. Levitt. 1971. Transformed Up-Down Methods in Psychoacoustics. *The Journal of the Acoustical Society of America* 49, 2B (Feb. 1971), 467–477. https://doi.org/10.1121/1.1912375

[33] Rémi Munos. 2011. Optimistic Optimization of a Deterministic Function without the Knowledge of its Smoothness. In *Advances in Neural Information Processing Systems 24*, J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 783–791. http://papers.nips.cc/paper/4304-optimistic-optimization-of-a-deterministic-function-without-the-knowledge-of-its-smoothness.pdf

[34] Rony-Reuven Nir, Yelena Granovsky, David Yarnitsky, Elliot Sprecher, and Michal Granot. 2011. A psychophysical study of endogenous analgesia: the role of the conditioning pain in the induction and magnitude of conditioned pain modulation. *European journal of pain* 15, 5 (2011), 491–497.

[35] Vishakha Patil, Ganesh Ghalme, Vineet Nair, and Y Narahari. 2020. Achieving Fairness in the Stochastic Multi-Armed Bandit Problem.. In *AAAI*. 5379–5386.

[36] Jonathan Peirce, Jeremy R. Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman, and Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51, 1 (Feb. 2019), 195–203. https://doi.org/10.3758/s13428-018-01193-y

[37] Jaelle Scheuerman, Kristen Brent Venable, Maxwell T Anderson, and Edward J Golob. 2017. Modeling Spatial Auditory Attention: Handling Equiprobable Attended Locations.. In *CAID@ IJCAI*. 1–7.

[38] Xuedong Shang, Emilie Kaufmann, and Michal Valko. 2019. General parallel optimization a without metric. In *Algorithmic Learning Theory*. 762–788. http://proceedings.mlr.press/v98/xuedong19a.html

[39] Yi Shen and Virginia M. Richards. 2012. A maximum-likelihood procedure for estimating psychometric functions: Thresholds, slopes, and lapses of attention. *The Journal of the Acoustical Society of America* 132, 2 (Aug. 2012), 957–967. https://doi.org/10.1121/1.4733540

[40] Aleksandrs Slivkins. 2014. Contextual bandits with similarity information. *The Journal of Machine Learning Research* 15, 1 (Jan. 2014), 2533–2568.

[41] Matthew Tesch, Jeff Schneider, and Howie Choset. 2013. Expensive Function Optimization with Stochastic Binary Outcomes. In *International Conference on Machine Learning*. 1283–1291. http://proceedings.mlr.press/v28/tesch13.html

[42] Rolf Ulrich and Jeff Miller. 2004. Threshold estimation in two-alternative forced-choice (2AFC) tasks: The Spearman-Kärber method. *Perception & Psychophysics* 66, 3 (April 2004), 517–533. https://doi.org/10.3758/BF03194898

[43] Michal Valko, Alexandra Carpentier, and Rémi Munos. 2013. Stochastic Simultaneous Optimistic Optimization. In *International Conference on Machine Learning*. 19–27. http://proceedings.mlr.press/v28/valko13.html

[44] William S. Verplanck, George H. Collier, and John W. Cotton. 1952. Nonindependence of successive responses in measurements of the visual threshold. *Journal of Experimental Psychology* 44, 4 (1952), 273–282. https://doi.org/10.1037/h0054948

[45] Felix A. Wichmann and N. Jeremy Hill. 2001. The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics* 63, 8 (Nov. 2001), 1293–1313. https://doi.org/10.3758/BF03194544

[46] Felix A. Wichmann and N. Jeremy Hill. 2001. The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics* 63, 8 (Nov. 2001), 1314–1329. https://doi.org/10.3758/BF03194545

[47] Yongwoo Yi and Daniel M. Merfeld. 2016. A quantitative confidence signal detection model: 1. Fitting psychometric functions. *Journal of Neurophysiology* 115, 4 (Jan. 2016), 1932–1945. https://doi.org/10.1152/jn.00318.2015

[48] Kamila Zchaluk and David H. Foster. 2009. Model-free estimation of the psychometric function. *Attention, Perception, & Psychophysics* 71, 6 (Aug. 2009), 1414–1425. https://doi.org/10.3758/APP.71.6.1414

[49] Tilmann Zwicker. 2000. Psychoacoustics as the basis for modern audio signal data compression. *The Journal of the Acoustical Society of America* 107, 5 (April 2000), 2875–2875. https://doi.org/10.1121/1.428677