STATISTICAL GUARANTEES IN THE SEARCH FOR LESS DISCRIMINATORY ALGORITHMS

Anonymous authors
Paper under double-blind review

ABSTRACT

Recent scholarship has argued that firms building data-driven decision systems in high-stakes domains like employment, credit, and housing should search for "less discriminatory algorithms" (LDAs) (Black et al., 2024). That is, for a given decision problem, firms considering deploying a model should make a good-faith effort to find equally performant models with lower disparate impact across social groups. Evidence from the literature on model multiplicity shows that randomness in training pipelines can lead to multiple models with the same performance, but meaningful variations in disparate impact. This suggests that developers can find LDAs simply by randomly retraining models. Firms cannot continue retraining forever, though, which raises the question: What constitutes a good-faith effort? In this paper, we formalize LDA search via model multiplicity as an optimal stopping problem, where a model developer with limited information wants to produce strong evidence that they have sufficiently explored the space of models. Our primary contribution is an adaptive stopping algorithm that yields a high-probability upper bound on the gains achievable from a continued search, allowing the developer to certify (e.g., to a court) that their search was sufficient. We provide a framework under which developers can impose stronger assumptions about the distribution of models, yielding correspondingly stronger bounds. We validate the method on real-world housing and lending datasets.

1 INTRODUCTION

Data-driven models increasingly underpin decision making in critical domains like employment, credit, and housing. While these models have been embraced for their potential to improve the quality and efficiency of such decision making, the literature on algorithmic fairness has shown that predictive models can also perpetuate or exacerbate societal biases, leading to potentially unfair outcomes (Barocas et al., 2023).

Recent work argues that in such high-stakes settings, firms building data-driven decision-making systems should proactively search for "less discriminatory algorithms" (LDAs) (Black et al., 2024; Gillis et al., 2024; Caro et al., 2024), or predictive models with equal overall performance but less "disparate impact" across legally protected groups. In the United States, disparate impact in these sectors is typically operationalized as the difference in selection rates across groups (e.g., differences in the hiring, lending, or renting rates across racial, gender, or age groups).

In support of their argument is the empirical finding that models optimized for accuracy can vary substantially with respect to other performance measures (like disparate impact), *even if the training procedure used is exactly the same* (Marx et al., 2020; D'Amour et al., 2022; Rudin et al., 2024; Black et al., 2022). This is because training processes are almost always non-deterministic; the subset of data used to train a model, the batch ordering in stochastic gradient descent, the set of features included as inputs, and any number of other aspects of a training algorithm are random. A firm might thus hope to sample a large set of models with comparable predictive performance and select the one with minimal disparate impact.

Scholars, advocates, and regulators have argued that firms are well-positioned to search for LDAs because they oversee model training (Black et al., 2024; FinRegLab, 2023; Blower, 2023). They have further argued that firms ought to take certain minimal steps to perform such searches, given that reductions in discrimination are sometimes achievable "for free" (i.e., without sacrificing accuracy) (Islam et al., 2021; Rodolfa et al., 2021). Others, however, have been more skeptical of the promise of LDAs, questioning whether they can really yield meaningful reductions in disparate impact and raising concerns about the lengths to which a firm must go to demonstrate a good-faith effort (Pace, 2023; Scherer et al., 2019). As one financial services blog put it: "no constraints or limits on this search have been proposed — and it is unclear how much resources, time, and effort are expected in searching for these potential LDAs" (Pace, 2022).

At the heart of this debate is the sense that a search for LDAs could potentially go on forever, given that additional searching might uncover an even less discriminatory alternative than what has been discovered already. Given this uncertainty, how can firms ever establish that they have performed a sufficient search for an LDA? In this paper, we develop a procedure for answering this question.

Our contributions. Our work develops statistical tools to quantify the value of an LDA search. We formalize LDA search as an optimal stopping problem, wherein a firm wants to continue training models as long as the marginal gain from doing so (in disparate impact reduction) is sufficiently large. Our primary contribution is an optimal stopping algorithm (Algorithm 1) and theorem (Theorem 3.5) to quantify and bound the value of continuing a search for LDAs. Our theorem provides a high-probability upper bound on the marginal value of training additional models, allowing a firm to stop an LDA search when its value is sufficiently low. Thus, our methods also provide a *certificate* of the limited benefits to a continued search, allowing the firm to demonstrate to a third party (e.g., a court or internal compliance team) that it has conducted a reasonable search. Our framework allows for the firm to impose knowledge about data and model distributions in order to further refine our algorithm's guarantees. Under stronger assumptions, we establish correspondingly stronger upper bounds on the marginal value of training additional models.

Beyond the LDA context, our algorithm establishes general high-probability guarantees for marginal returns of additional samples when sampling from an unknown distribution. At a technical level, we draw on recent results on *anytime-valid inference*, which allow us to adaptively stop training models while maintaining statistical validity. In particular, we develop a novel and asymptotically near-optimal sequence upper-bounding the probability of improving upon a running best sample drawn iid from any distribution, which may be of independent interest.

We also evaluate our algorithm empirically on a number of publicly available datasets related to credit and housing. We randomly retrain models across standard model classes and measure the stopping time of our algorithm against the optimal full-information stopping time. We find significant heterogeneity in the true, full-information marginal returns to retraining, and in the performance of the algorithm relative to this idealized benchmark.

Related work. Our technical approach is most closely related to a long literature in economics and computer science on optimal stopping (DeGroot, 2004; Beyhaghi & Cai, 2024; Lippman & McCall, 1976; Bikhchandani & Sharma, 1996). Our model is closely related to the Pandora's Box problem (Weitzman, 1978; Kleinberg et al., 2016; Beyhaghi & Cai, 2024), in which the decision-maker pays a cost to sample from a *known* distribution. However, our work is different in that we assume minimal knowledge of the distributions. Also, rather than trying to maximize total utility of a search, we seek a high-probability guarantee on the marginal returns of drawing another sample.

Second, our work is motivated by a literature on less discriminatory algorithms, model multiplicity and fairness/accuracy tradeoffs (Black et al., 2024; 2022; Rodolfa et al., 2021; Laufer et al., 2025; Gillis et al., 2024; Cen et al., 2025; Fallah et al., 2025; Rudin et al., 2024). This literature surfaces the idea that there may be many highly accurate models, and that retraining models may yield predictors with different properties, especially with respect to fairness. Our work addresses an important and unanswered question in this area: *How do we certify the sufficiency of a search for a particular model retraining process?*

Organization of the paper. In Section 2, we formalize our setting, including the model retraining process and our goals. In Section 3, we describe our theoretical results, including an algorithm for adaptively training models and a theorem with corresponding guarantees on the correctness of the stopping time. In Section 4, we validate our method on real-world datasets for credit and housing. Finally, in Section 5, we discuss other applications of our technical approach and conclude.

2 Setting and model

At a high level, we study the problem of learning a predictive machine learning model from a finite dataset. The firm's utility for a predictor is determined by its average performance on a loss function over the population distribution. In the search for LDAs, the loss function might be the difference in selection rates of a protected group versus that of a reference group.

The firm seeks to take advantage of model multiplicity to reduce this loss by sampling multiple high performing models and selecting the least discriminatory among them (i.e., the one that minimizes disparate impact). Our target is to design a procedure which determines when a sufficient search has been conducted during the re-training process.

We assume the model trainer pre-specifies (1) a cost for sampling an additional model by repeating a randomized training procedure and (2) a utility for a unit improvement to disparate impact. The ratio of these quantities specifies a target threshold for determining whether the marginal benefit of retraining models is worth the cost: if the expected benefit from training a new model is above the threshold, the model trainer should do so, and if it is below the threshold, the trainer should terminate the retraining procedure and deploy the best model seen so far. In the remainder of this section, we formalize this setting and define notation.

Data and utility. We will assume the existence of an unknown population distribution $\mathcal{D}$ from which the firm has sampled an iid dataset D of size n, consisting of labeled data pairs $(x, y) \in \mathcal{X} \times \mathcal{Y}$. The firm will deploy a predictor $h : \mathcal{X} \to \mathcal{Y}$. In cases where the predictor determines outcomes (like offers of employment, credit, or housing), $\mathcal{Y}$ will be binary, where 1 is the positive outcome. The firm's utility will be defined as

$$Q(h) \triangleq \mathbb{E}_{(x,y)\sim\mathcal{D}}\left[\ell(h(x), y, x)\right]$$

for ℓ given and $\mathrm{im}(Q) \subseteq [0,1]$.[1] If the goal is to reduce disparity in selection rates with respect to a group indicator $g(x) \in \{0,1\}$ as in the search for LDAs, ℓ would be written as $\ell^{\mathrm{DI}}(a, y, x) = \left((1-g(x))/P(g(x)=0) - g(x)/P(g(x)=1)\right) a$, which is $a/P(g(x)=0)$ if $g(x)=0$ and $-a/P(g(x)=1)$ if $g(x)=1$. Then, the expected selection rate disparity of the model would be given by $Q^{\mathrm{DI}}(h) = \mathbb{E}[h(X) \mid g(X)=0] - \mathbb{E}[h(X) \mid g(X)=1]$, i.e., the difference between the selection rate for the reference group and the selection rate for the protected group. The loss is bounded in $[0,1]$ if the selection rate for the protected group (X for which $g(X)=1$) is never greater than that for the reference group (X for which $g(X)=0$). This is reasonable because discrimination against the protected group is not a concern if their selection rate is higher than the reference group.[2] However, our results are not solely relevant to ℓ as the selection rate disparity: our results hold for any outcome space $\mathcal{Y}$ and loss function ℓ as long as the range of Q is bounded in $[0,1]$.

The model trainer cannot observe their true utility. Instead, we will assume they have access to a finite sample of data on which they will evaluate their model. The empirical performance will be defined for a fixed dataset S, as

$$\hat{Q}(h; S) \triangleq \frac{1}{|S|}\sum_{i\in S} \ell(h(x_i), y_i, x_i).$$

Model distribution. The model trainer will have a randomized training procedure $\mathcal{A}$ that takes in a dataset D and returns a model h. There are no assumptions on the procedure $\mathcal{A}(D)$, except that it is fixed in advance and returns a model iid conditional on the data D.

While we assume the model trainer has a fixed dataset D, we do not necessarily assume that all models are trained on the same training sample. Instead, the data may be partitioned into subsets D^{train} and D^{test}, where $h = \mathcal{A}(D)$ depends only on the training subset D^{train} and not on the remaining data $D^{\mathrm{test}} = D \backslash D^{\mathrm{train}}$. Additionally, $\mathcal{A}$ is not restricted to produce models from any particular

[1] This is without loss of generality: any bounded loss function can be rescaled so the loss is on $[0,1]$. Our proposed methods therefore work for loss functions beyond disparate impact; for example, a firm could minimize a weighted combination of disparate impact and error rate instead of disparate impact alone.

[2] If selection rate disparity is a concern for both groups (i.e., both groups are protected), this loss could alternately represent the absolute value of the difference between selection rates between groups.

model class or setting of hyperparameters—it does not need to be a standard model training process for a fixed model class. For example, $\mathcal{A}$ might first randomly decide between multiple algorithms (which themselves might be randomized), like random forests or neural networks. Alternately, $\mathcal{A}$ might sample from a given distribution over hyperparameters.

We will analyze the setting in which a model trainer trains a sequence of models $h_1, h_2, \ldots$ by sampling iid, conditional on D, from $\mathcal{A}(D)$. Let $D_1^{\text{train}}, D_2^{\text{train}}, \ldots$ be the sequence of training splits and $D_1^{\text{test}}, D_2^{\text{test}}, \ldots$ be the sequence of test splits. (Recall $D_t^{\text{train}} \cup D_t^{\text{test}} = D$ for all t, so train and test splits for different steps t will have shared data.) For brevity, we will write the true and empirical loss of the t-th model as

$$Q_t \triangleq Q(h_t), \qquad \text{and} \qquad \hat{Q}_t \triangleq \hat{Q}(h_t; D_t^{\text{test}}).$$

We will denote by P the distribution of the infinite sequence $Q_1, Q_2, \ldots$. (When just considering the first t entries of this sequence, we will imagine throwing the rest away so as to not introduce new notation.) We will denote by $\hat{P}$ the distribution of the infinite sequence $\hat{Q}_1, \hat{Q}_2, \ldots$ similar to P, and assume that P and $\hat{P}$ are defined on the same space. All distributions and probabilities throughout this work are taken conditional on D, since we imagine there is one fixed dataset used for training and evaluation. Note that P and $\hat{P}$ are supported on (a subset of) $[0, 1]^\infty$, since $Q_t, \hat{Q}_t \in [0, 1]$ by assumption. Also, note that $\{Q_t\}_{t=1}^\infty$ are iid, conditional on D. Let P_0 be the marginal distribution of any Q_t. Similarly, $\{\hat{Q}_t\}_{t=1}^\infty$ are iid conditional on D and we will denote the marginal distribution of any $\hat{Q}_t$ by $\hat{P}_0$. Finally, let $\mathbb{P}$ be the joint probability distribution over the pairs $(Q_1, \hat{Q}_1), \ldots$ and let $\mathbb{P}_0$ be the marginal distribution over any $(Q_t, \hat{Q}_t)$.

We will analyze the model with the best performance on the test split, after the trainer concludes training. Formally, for given t, let i_t be the model with the lowest empirical disparate impact up to the t-th model: $i_t = \arg\min_{i \in [t]} \hat{Q}_i$. We will analyze the case where, after the model trainer trains τ models, they select and deploy h_{i_τ}. The true and empirical disparate impact of the *selected* model after training t models will be denoted

$$U_t \triangleq Q_{i_t}, \qquad \text{and} \qquad \hat{U}_t \triangleq \hat{Q}_{i_t}.$$

In the context of an LDA search, we assume the models sampled from $\mathcal{A}(D)$ are all *deployable*, in the sense that a sample from $\mathcal{A}(D)$ meets the business needs of the firm. If this is not true for some model training process, rejection sampling can be used to continue retraining until a deployable one is found. In practice, this may be accomplished by, for example, setting an accuracy threshold and letting $\mathcal{A}(D)$ be samples from the model training distribution, conditional on sufficient accuracy.[3]

Certifying a sufficient search. For given cost of training a single model c and utility for a unit improvement to disparate impact b, the model trainer is justified in terminating a search after training τ models if $b \cdot \mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \le c$, i.e., the expected marginal benefit of training an additional model, given the observed best model so far, does not outweigh the cost. Equivalently, we will write

$$\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \le \gamma \tag{1}$$

where we define $\gamma \triangleq c/b$. Our definition requires the model trainer to continue sampling models as long as the expected benefits outweigh the cost. But our information is limited in two ways: First, we do not know P. Second, we can only observe noisy estimates of Q_t due to our finite data sample. Thus, we can only hope to upper bound the left-hand expression of eq. (1), with high probability over τ, given this uncertainty.

3 Adaptive Stopping for Repeated Model Retraining

Our main theoretical contribution is an adaptive algorithm (Algorithm 1) and accompanying theoretical result (Theorem 3.5). The algorithm gives a procedure for training models until a stopping

[3] In other settings, ℓ might represent accuracy itself, in which case the search would be for more accurate models. However, our motivation for this work is clarifying the debate around LDAs. The model multiplicity literature argues that models optimized for accuracy will have similar accuracy but perhaps differences in other properties (Black et al., 2022; Rodolfa et al., 2021). More generally, ℓ could encode some fairness-accuracy trade-off via a weighted combination of different objectives.

condition is met. The theorem establishes that, when the algorithm halts, the marginal benefits of retraining can be concluded to be no longer worth the costs. We also establish that the algorithm always halts at some finite time that depends on γ and gives a data-independent upper bound on the number of models that need to be trained.

Our plan for the section is as follows. To build intuition, in Sections 3.1 and 3.2, we start with analyses of simpler settings. In Section 3.1, the distribution of model performance is known, and observations of performance are observed exactly as if they were evaluated on infinite data (i.e., $\hat{Q}_t = Q_t$ for all t). In this regime, the stopping problem is trivial and can be described by a threshold on draws from the model peformance distribution. Next, in Section 3.2, we relax the first condition and do not assume full knowledge of the model performance distribution. We outline how different conditions on the model performance distribution yield different bounds, and our method allows decision-makers to input assumptions suitable to their context. Then, in Section 3.3, we handle the additional uncertainty from evaluations on finite data. To do so, we introduce a natural assumption on the relationship between observed and true model performance. Finally, in Section 3.4, we consider the case in which estimation of a property of the model loss distribution can be leveraged to produce tighter bounds on marginal benefits of model retraining. All proofs are deferred to Appendix D.

3.1 The full-information regime

We first consider the simplest case, when both the distribution P is known and the population values of Q_t are exactly observed. For any t, note that $U_t - U_{t+1} = (U_t - Q_{t+1}) \cdot \mathbb{I}[U_t > Q_{t+1}]$. Thus, if the performance of the best model so far is u, the expected marginal gain of a new sample is

$$g(u) \triangleq \mathbb{E}_{Q\sim P_0}\left[(u - Q) \cdot \mathbb{I}\left[u > Q\right]\right]. \tag{2}$$

Observe that g is weakly monotonically increasing, and $g(0) = 0$. Therefore, there is some threshold u_P^* at which the marginal gain drops below γ. Define this threshold as follows:

$$u_P^* \triangleq \sup_{u\in[0,1]} \{u : g(u) \leq \gamma\}.$$

Thus, our stopping time τ satisfies the desired guarantee eq. (1) if and only if

$$\mathbb{E}_{P_0}\left[U_\tau - U_{\tau+1} \mid U_\tau\right] \leq \gamma \iff g(U_\tau) \leq \gamma \iff U_\tau \leq u_P^*. \tag{3}$$

This immediately yields a stopping condition: compute u_P^* and sample until a value less than u_P^* is observed. The stopping time τ in this case is geometrically distributed, since each sample is less than u_P^* with probability $P_0(u_P^* \geq U_{\tau+1})$, and so the expected stopping time is $1/P_0(u_P^* \geq U_{\tau+1})$.

3.2 The infinite-data regime

Next, we analyze the case where we can perfectly observe Q_t for all t. In this case, our only source of uncertainty is our lack of information about P. Because of our uncertainty about P, we cannot always guarantee eq. (1) for finite τ: there is always a chance that the sequence $\{Q_s\}_{s=1}^t$ observed so far have been abnormally large (i.e., an especially unlucky sequence), so that the expected marginal gain of a new sample is greater than γ. The best we can do is ensure that it holds *with high probability*, over the randomness of $\{Q_t\}_{t=1}^\infty$. That is, for a pre-specified $\delta \in (0, 1)$, we want

$$P(\mathbb{E}_{P_0}\left[U_\tau - U_{\tau+1} \mid U_\tau\right] \leq \gamma) = P(g(U_\tau) \leq \gamma) \geq 1 - \delta. \tag{4}$$

where the expectation is over $U_{\tau+1}$ marginally and the probability is over all t jointly. Our goal is thus to provide an anytime-valid upper bound on $\{g(U_t)\}_{t=1}^\infty$. That is, suppose we had a sequence $\{\bar{g}_t(U_t)\}_{t=1}^\infty$ such that

$$P(\exists t \in \mathbb{N} \ : \ g(U_t) > \bar{g}_t(U_t)) \leq \delta.$$

Then, it suffices to stop sampling at τ such that $\bar{g}_\tau(U_\tau) \leq \gamma$.

We have thus reduced our stopping problem to maintaining an anytime-valid upper bound for $g(U_t)$. Our next step is to actually construct such a bound. To do so, we decompose $g(\cdot)$ into two terms: One which captures the probability of observing a strictly better sample, and another which captures the expected improvement *conditional* on observing a strictly better sample. Observe

$$g(u) = \mathbb{E}_{Q\sim P_0}\left[u - Q \mid u > Q\right] P_0(u > Q) = \mu(u)p(u),$$

Assumption	Interpretation	$\bar{\mu}$
No assumption	Applies to any distribution	$\bar{\mu}^{\text{universal}}(u) \triangleq u$
(A1) $\exists a > 0$ s.t. $f_{P_0}(x)$ is increasing for $x \le a$	P_0 has a sub-uniform left tail	$\bar{\mu}^{\text{mono}}(u) \triangleq \begin{cases} u & u > a \\ \frac{u}{2} & u \le a \end{cases}$
(A2) $\exists a > 0$ s.t. $\mu(u)$ is increasing for $u \le a$	P_0 has an exponential or sharper left tail	$\bar{\mu}^{\text{exp}} \triangleq \begin{cases} u & u > a \\ \min(\mu(a), u) & u \le a \end{cases}$
(A3) $\exists a > 0$ s.t. $P_0(Q < a) = 0$	No model has disparate impact lower than a	$\bar{\mu}^{\text{bounded}} \triangleq u - a$

Table 1: Assumptions on P_0 and corresponding bounds $\bar{\mu}$.

where we define, for a draw of Q iid from P_0, $\mu(u) \triangleq \mathbb{E}_{P_0}[u - Q \mid u > Q]$ and $p(u) \triangleq P_0(u > Q)$. We will call μ the *conditional expected improvement* (CEI)[4] and p the *improvement probability*. It suffices to upper bound each of these separately and then combine them.

Bounding μ. We first formalize a definition for bounds on μ which will allow us to plug in different bounds for different conditions on the input distribution. The definition is written for a generic distribution since we will reuse this definition later in the finite-data case.

Definition 3.1 ($\bar{\mu}$-Bounded CEI for $\mathcal{P}$)**.** $\bar{\mu} : [0,1] \to [0,1]$ is a CEI bound for distribution $\mathcal{P}$ if

$$\mathbb{E}_{Q\sim\mathcal{P}}[u - Q \mid u > Q] \le \bar{\mu}(u)$$

for all $u \in [0,1]$, almost surely.

Next, we provide a series of assumptions on P_0 under which we can derive bounds $\bar{\mu}$ satisfying Definition 3.1. These are summarized in Table 1. First, note that $\mu(U_t) \le U_t$ almost surely since $U_{t+1} \ge 0$. This bound is quite conservative, since it bounds *expected improvement* by *maximum possible improvement.* Decision-makers can make stronger assumptions on P_0 to get tighter bounds. For example, consider the case when P_0 is continuous, and there exists $a \in (0,1)$ such that $f_{P_0}(x)$ is non-decreasing in x for all $x \le a$ (i.e., P_0, at worst, has a uniform-like left tail). Note that, by this assumption, $\mu(u) \le \int_0^u x/u \, \mathrm{d}x = u/2$. Thus, we can define $\bar{\mu}^{\text{mono}}(u)$ as $u/2$ if $u \le a$ and u otherwise. Similarly, if there exists some $a \in (0,1)$ such that the $\mu(u)$ is increasing for all $u \le a$, then we can apply $\bar{\mu}^{\text{exp}}(u) = \min(\mu(a), u)$ if $u \le a$ and u otherwise. Finally, suppose $P_0(Q < a) = 0$ for some $a \in (0,1]$. Then we can define $\bar{\mu}_t^{\text{bounded}}(u) \triangleq u - a$.

Bounding p. The following lemma yields a general anytime-valid high probability upper bound for the probability of observing a new minimum in a sequence of iid random variables. It may be of independent interest. An asymptotically near-optimal (but more complex) sequence can be found in Theorem E.1.

Lemma 3.2. *Let $\{X_t\}_{t=1}^\infty$ be a sequence of iid random variables distributed according to a law $\mathcal{P}_0$. Let $\mathcal{P} \triangleq \mathcal{P}_0^\infty$ be their joint distribution. Let $Y_t \triangleq \min_{s\in[t]} X_s$. For any $\alpha \in (0,1)$, define*

$$\bar{p}_t(\alpha) = \begin{cases} 1 - e^{-1/\alpha} & \text{if } t = 1 \\ 1 - \left(\frac{(t-1)}{\alpha} + 1\right)^{-1/(t-1)} & \text{otherwise.} \end{cases}$$

Then,

$$\mathcal{P}(\exists t \in \mathbb{N} \; : \; \mathcal{P}_0(X_{t+1} < Y_t \mid Y_t) > \bar{p}_t(\alpha)) \le \alpha.$$

Lemma 3.2 yields an immediate anytime-valid upper bound on $\{p(U_t)\}$:

$$P(\exists t \in \mathbb{N} \; : \; p(U_t) > \bar{p}_t(\delta)) \le \delta. \tag{5}$$

[4]This concept is closely related to that of the *mean residual life* of a random variable, for which there is a rich literature. See, e.g., Hall & Wellner (2020).

Combining bounds. Our algorithm simply combines our bounds on μ and p to maintain an anytime-valid upper bound on the marginal gain, given by $\bar{\mu}(U_t) \cdot \bar{p}_t(\delta)$. Formally, our algorithm simply terminates at the first τ such that $\bar{\mu}_\tau(U_\tau) \cdot \bar{p}_\tau(\delta) \leq \gamma$. Moreover, τ is guaranteed to be finite because $\bar{\mu}_t(\cdot) \leq 1$ for all t, and $\lim_{t\to\infty} \bar{p}_t(\delta) = 0$. A data-independent upper bound on the number of models trained can thus be directly computed from δ and γ by finding the t such that $p_t(\delta) < \gamma$. We state the algorithm for a generic distribution $\mathcal{P}$ given as input, rather than P_0, since we will reuse this algorithm in the finite-data regime.

Algorithm 1 LDA Search with Adaptive Stopping

---

input:
  An unknown model performance distribution $\mathcal{P}$ from which to draw iid samples.
  Stopping threshold γ and failure probability δ.
  Optional: An almost-sure expected conditional improvement bound $\bar{\mu}$ satisfying Definition 3.1.
  If not provided, use $\bar{\mu}^{\text{universal}}(u) = u$.

1: **for** $t = 1, 2, \ldots$ **do**
2: Draw a new sample $X_t \overset{\text{iid}}{\sim} \mathcal{P}$.
3: Define $\bar{p}_t$ as in Lemma 3.2.
4: Define $Y_t = \min_{s \leq t} X_t$
5: **if** $\bar{\mu}(Y_t) \cdot \bar{p}_t(\delta) < \gamma$ **then**
6: **return** Y_t
7: **end if**
8: **end for**

We now state the formal statistical guarantee for our infinite data setting. It is a special case of a more general theorem we prove, Theorem D.1.

Proposition 3.3. *For all $\gamma, \delta > 0$, Algorithm 1 run with $\mathcal{P} = P_0$, γ, δ and any $\bar{\mu}$ that satisfies Definition 3.1 for P_0 as input terminates at a stopping time $\tau \in \mathbb{N}$ such that*

$$P(\mathbb{E}_{P_0}[U_\tau - U_{\tau+1} \mid U_\tau] < \gamma) \geq 1 - \delta.$$

Next, we generalize to the case where we have finite data.

3.3 The finite-data regime

If we observe only finite data, we cannot perfectly observe each Q_t; instead, we observe $\hat{Q}_t$. As before, we will seek to maintain an anytime-valid upper bound on the marginal gain. We must take care to define the marginal gain appropriately—in particular, our goal is to bound the expected marginal gain with respect to the *true* disparate impact (Q_t), given our observations of empirical disparate impact ($\hat{Q}_t$). Formally, our goal is to show that, at stopping time τ,

$$\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma.$$

where the expectation is also conditional on D. To do this, we need to establish a relation between the measurement error $U_t - \hat{U}_t$ at different points on the left tail of $\hat{P}_0$. We provide a natural assumption on the relationship between these quantities: the selection effect or regression-to-the-mean effect is, in expectation, non-decreasing in t. The assumption that regression-to-the-mean is at least constant is frequently supposed in the large literature on adjusting analysis for or estimating these effects (Stein et al., 1956; James et al., 1961; Sorensen & Kennedy, 1984; Andrews et al., 2024; Zrnic & Fithian, 2024; Fithian et al., 2014). Intuitively, this assumption holds for sub-Gaussian left tails where the selection effect should be linear in the gap between $\hat{U}_t$ and $\hat{U}_{t+1}$ and even for sub-exponential left tails where there should be constant regression to the mean in the gap between $\hat{U}_t$ and $\hat{U}_{t+1}$. This assumption would not hold if some measurable set of values of $\hat{U}_t$ indicate that the model is extremely fair, while models with $\hat{U}_{t+1} < \hat{U}_t$ are not particularly fair.

Assumption 3.4 (Non-decreasing selection effect)**.** It holds for all t that

$$\mathbb{E}_{\mathbb{P}_0}[U_t - \hat{U}_t \mid \hat{U}_t] \geq \mathbb{E}_{\mathbb{P}_0}[U_{t+1} - \hat{U}_{t+1} \mid \hat{U}_t].$$

Under Assumption 3.4, we can apply the same algorithm on the sequence $\{\hat{U}_t\}_{t=1}^{\infty}$ that we applied to $\{U_t\}_{t=1}^{\infty}$ in the infinite data case. This additional assumption is sufficient for the following theorem to hold, using only a minor modification to the argument applied in the infinite data case.

Theorem 3.5. *Under Assumption 3.4, for all $\gamma > 0$ and $\delta > 0$, Algorithm 1 run with $\mathcal{P} = \hat{P}_0$, γ, δ and any $\bar{\mu}$ that satisfies Definition 3.1 for $\hat{P}_0$ terminates at a time $\tau \in \mathbb{N}$ such that*

$$\mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) \geq 1 - \delta. \tag{6}$$

3.4 Data-driven anytime-valid upper bounds on the conditional expected improvement

We conclude this section with a discussion of the case in which $\bar{\mu}$ can be estimated from data. We provide intuition here and defer formal analysis to Appendix C.

A model developer may reasonably believe that Assumption (A2) holds, meaning the conditional expected improvement is *decreasing* as we sweep towards the left tail of $\hat{P}_0$. However, they may have no *a priori* knowledge of the precise value of that bound, given by $\mu(a)$. The developer can instead infer a high-probability anytime-valid upper bound on $\mu(a)$ under Assumption (A2), yielding a data-driven CEI bound $\bar{\mu}^{\text{exp}}$. In Appendix C, we provide an algorithm (Algorithm 2) to formalize this idea, taking care to combine multiple anytime-valid bounds.

4 Empirical Analysis

In this section, we evaluate our method on several datasets and model classes. The datasets we use are Adult (Becker & Kohavi, 1996), Folktables (Ding et al., 2021), and HMDA (CFPB, 2017). The first two of these are lending prediction tasks and the third is a mortgage prediction task. The model classes we use are logistic regression, random forests and neural networks.

To evaluate our algorithm, we would ideally compare the performance of our algorithm against the full-information regime discussed in Section 3.1, where we perfectly observe the marginal benefit of sampling a new model. This is in general not possible, since we know neither the true data distribution nor the true distribution of model disparate impacts. Instead, we treat the finite dataset as a "population" and subsample to produce semi-synthetic datasets. We use a similar technique to subsample from a large pool of trained models. Further details on our data preparation, model training and comparison to the full-information regime are available in Appendix B.

The results of running the algorithm on many subsamples are visualized in Figure 1. Iterations of the algorithm are on the horizontal axis and the conditional expected improvement is on the vertical axis. The pink line is our upper bound $\bar{\mu}(\hat{U}_t)\bar{p}_t(\delta)$ for $\delta = 0.05$, setting $\bar{\mu}(\hat{U}_t) = \hat{U}_t$, meaning we place no assumptions on the distribution $\hat{P}_0$. The brown line is the ground truth $g(\hat{U}_t)$. The shaded colored regions for each line show standard deviations over multiple runs of the dataset resampling, model training and algorithm.

For any γ, Algorithm 1 would stop when the pink line drops below the horizontal line at γ. Given full distributional information, a model trainer should stop the retraining process once the brown line drops below the horizontal line at γ. Thus, for any fixed γ, the average number of iterations that the algorithm trained models past the stopping time given full information is the horizontal distance between the brown and pink lines. Empirically, Algorithm 1 performs well in the sense that it "overshoots" the correct stopping time by tens of models in general, though it appears to perform worse for logistic regression. Further assumptions (i.e., A1, A2, A3) will likely yield tighter bounds. We provide miscoverage rates for our upper bound in Figure 4, which are well below the target coverage 0.05 on average across datasets and model classes.

5 Discussion

Although recent work has proposed that firms should take steps to proactively search for less discriminatory algorithms, there are a number of open questions regarding both the gains to be expected from an LDA search and the resources required to conduct one. In this paper we take one step towards developing the tooling firms would need to conduct a search. We put forward a method that

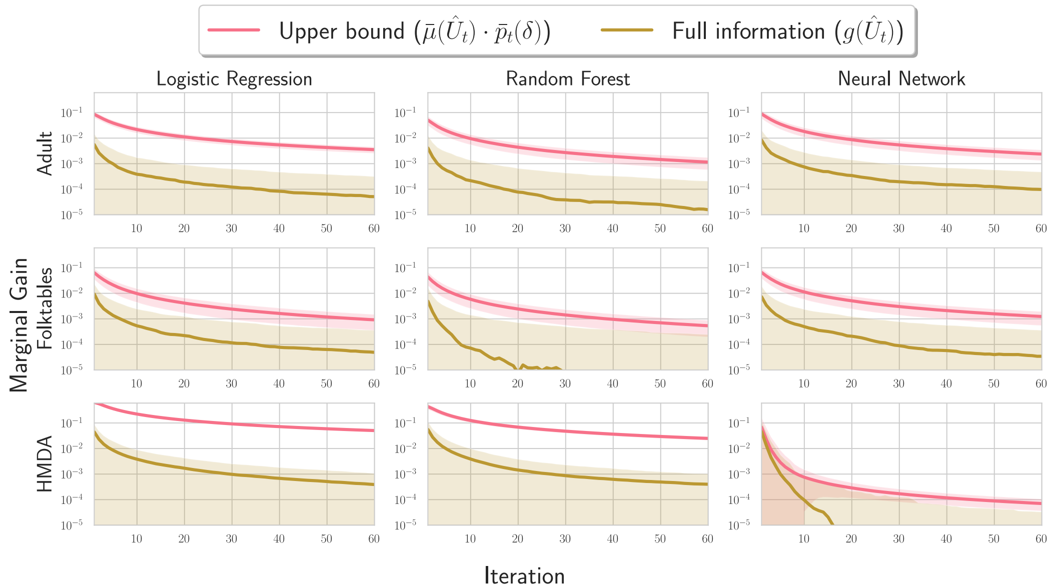

Figure 1: Algorithm 1 run on several datasets and models. Panel rows are datasets and panel columns are model classes. In each panel, the horizontal axis is the iteration of the algorithm and the vertical axis is marginal gain. The pink line is our estimated upper bound $\bar{\mu}(\hat{U}_t)\bar{p}_t(0.05)$ and the brown line is the full-information marginal gain. For any γ, Algorithm 1 would stop when the pink line crosses the horizontal line at γ. Note that the vertical axis is on a log scale.

allows firms to adaptively sample models that come from a particular loss distribution. Our algorithm adaptively bounds the marginal gains of a continued search, allowing a firm to terminate the search when the gains are small and provide evidence that their search was sufficient.

We take as given γ, which specifies the developer's cost of training models relative to their value of reducing disparate impact. While determining how a firm might choose γ is beyond the scope of this work, it is the subject of ongoing debate (Pace, 2022; Black et al., 2024). Our framework can help contribute to this debate in at least two ways. First, because we provide anytime-valid bounds, we do not require that a firm pre-specify γ. Instead, model developers and compliance teams can iteratively develop models, consider the incremental gains, run separate experiments, and adaptively decide how to value those gains relative to development costs. Second, given a search conducted by a firm, our framework allows us to "back out" a high-probability upper bound on the firm's value of γ implied by their decision to stop the search. That is, by observing a sequence of models sampled by a developer, we can draw conclusions about their implicit value for reducing disparate impact from their decision to terminate a search, and thereby facilitate a more informed debate about the reasonableness of the search.

Our proposed procedure is just one piece of a larger and more complex set of steps that a firm might take to search for a less discriminatory algorithm. This should not be construed as the only thing that the firm has to do. In real cases, debates about the existence of a less discriminatory algorithm might cover a swath of both quantitative and qualitative considerations about the reasonableness of model assumptions, variables used, and so forth (Black et al., 2024).

A number of future directions related to this setting are open. Our framework could be extended to handle adaptivity, where the performance of previous models informs training decisions for future models. In low-data settings, we would expect *shrinkage* or *selection effects* to be salient: the best-performing model in-sample could fare much worse out-of-sample, potentially admitting stronger guarantees. Finally, our technical framework can be applied to general optimal stopping problems where high-probability guarantees are desirable. For example, a developer or researcher using an LLM may choose the best of many randomly sampled prompts, and with our algorithm, they can certify that further exploration is unlikely to yield significant gains. Applying our framework to other settings is a fruitful direction for future work.

REFERENCES

Isaiah Andrews, Toru Kitagawa, and Adam McCloskey. Inference on Winners. *The Quarterly Journal of Economics*, 139(1):305–358, January 2024. ISSN 0033-5533, 1531-4650. doi: 10.1093/qje/qjad043. URL https://academic.oup.com/qje/article/139/1/305/7276491.

Solon Barocas, Moritz Hardt, and Arvind Narayanan. *Fairness and machine learning: Limitations and opportunities*. MIT press, 2023.

Barry Becker and Ronny Kohavi. Adult. UCI Machine Learning Repository, 1996. DOI: https://doi.org/10.24432/C5XW20.

Hedyeh Beyhaghi and Linda Cai. Recent developments in pandora's box problem: Variants and applications. *ACM SIGecom Exchanges*, 21(1):20–34, 2024.

Sushil Bikhchandani and Sunil Sharma. Optimal search with learning. *Journal of Economic Dynamics and Control*, 20(1-3):333–359, 1996.

Emily Black, Manish Raghavan, and Solon Barocas. Model multiplicity: Opportunities, concerns, and solutions. In *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 850–863, 2022.

Emily Black, John Logan Koepke, Pauline T Kim, Solon Barocas, and Mingwei Hsu. Less discriminatory algorithms. *Geo. LJ*, 113:53, 2024.

Brad Blower. CFPB Puts Lenders & FinTechs On Notice: Their Models Must Search For Less Discriminatory Alternatives Or Face Fair Lending Non-Compliance Risk NCRC, April 2023. URL https://www.ncrc.org/cfpb-puts-lenders-fintechs-on-notice-their-models-must-search-for-less-discriminat Section: Views.

Spencer Caro, Talia B Gillis, and Scott Nelson. Modernizing fair lending. *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2024-18), 2024.

Sarah H. Cen, Salil Goyal, Zaynah Javed, Ananya Karthik, Percy Liang, and Daniel E. Ho. Audits Under Resource, Data, and Access Constraints: Scaling Laws For Less Discriminatory Alternatives, September 2025. URL http://arxiv.org/abs/2509.05627. arXiv:2509.05627 [cs].

CFPB. Home mortgage disclosure act data. United States Government, 2017. URL https://www.consumerfinance.gov/data-research/hmda/historic-data/.

A Feder Cooper, Katherine Lee, Madiha Zahrah Choksi, Solon Barocas, Christopher De Sa, James Grimmelmann, Jon Kleinberg, Siddhartha Sen, and Baobao Zhang. Arbitrariness and social prediction: The confounding role of variance in fair classification. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 22004–22012, 2024.

Alexander D'Amour, Katherine Heller, Dan Moldovan, Ben Adlam, Babak Alipanahi, Alex Beutel, Christina Chen, Jonathan Deaton, Jacob Eisenstein, Matthew D Hoffman, et al. Underspecification presents challenges for credibility in modern machine learning. *Journal of Machine Learning Research*, 23(226):1–61, 2022.

Morris H. DeGroot. *Optimal Stopping*, pp. 324–384. John Wiley & Sons, Ltd, 2004. ISBN 9780471729006. doi: https://doi.org/10.1002/0471729000.ch13. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/0471729000.ch13.

Frances Ding, Moritz Hardt, John Miller, and Ludwig Schmidt. Retiring adult: New datasets for fair machine learning. *Advances in Neural Information Processing Systems*, 34, 2021.

Alireza Fallah, Michael I. Jordan, and Annie Ulichney. The Statistical Fairness-Accuracy Frontier, August 2025. URL http://arxiv.org/abs/2508.17622. arXiv:2508.17622 [stat].

FinRegLab. Machine learning explainability & fairness: Insights from consumer lending. Empirical white paper, FinRegLab, July 2023. URL https://finreglab.org/wp-content/uploads/2023/12/FinRegLab_2023-07-13_Empirical-White-Paper_Explainability-and-Fairness_Insights-from-Consumer-Lending.pdf. Updated July 2023.

William Fithian, Dennis Sun, and Jonathan Taylor. Optimal inference after model selection. *arXiv preprint arXiv:1410.2597*, 2014.

Talia B Gillis, Vitaly Meursault, and Berk Ustun. Operationalizing the Search for Less Discriminatory Alternatives in Fair Lending. In *The 2024 ACM Conference on Fairness, Accountability, and Transparency*, pp. 377–387, Rio de Janeiro Brazil, June 2024. ACM. ISBN 979-8-4007-0450-5. doi: 10.1145/3630106.3658912. URL https://dl.acm.org/doi/10.1145/3630106.3658912.

W. J. Hall and Jon A. Wellner. Estimation of Mean Residual Life. In Anthony Almudevar, David Oakes, and Jack Hall (eds.), *Statistical Modeling for Biological Systems*, pp. 169–189. Springer International Publishing, Cham, 2020. ISBN 978-3-030-34674-4 978-3-030-34675-1. doi: 10.1007/978-3-030-34675-1_10. URL http://link.springer.com/10.1007/978-3-030-34675-1_10.

Rashidul Islam, Shimei Pan, and James R. Foulds. Can We Obtain Fairness For Free? In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 586–596, Virtual Event USA, July 2021. ACM. ISBN 978-1-4503-8473-5. doi: 10.1145/3461702.3462614. URL https://dl.acm.org/doi/10.1145/3461702.3462614.

William James, Charles Stein, et al. Estimation with quadratic loss. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 1, pp. 361–379. University of California Press, 1961.

Robert Kleinberg, Bo Waggoner, and E. Glen Weyl. Descending price optimally coordinates search. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, EC '16, pp. 23–24, New York, NY, USA, 2016. Association for Computing Machinery. ISBN 9781450339360. doi: 10.1145/2940716.2940760. URL https://doi.org/10.1145/2940716.2940760.

Benjamin Laufer, Manish Raghavan, and Solon Barocas. What Constitutes a Less Discriminatory Algorithm? In *Proceedings of the Symposium on Computer Science and Law on ZZZ*, pp. 136–151, March 2025. doi: 10.1145/3709025.3712214. URL http://arxiv.org/abs/2412.18138. arXiv:2412.18138 [cs].

Steven A Lippman and John J McCall. The economics of job search: A survey. *Economic inquiry*, 14(2):155–189, 1976.

Charles Marx, Flavio Calmon, and Berk Ustun. Predictive Multiplicity in Classification. In *Proceedings of the 37th International Conference on Machine Learning*, pp. 6765–6774. PMLR, November 2020. URL https://proceedings.mlr.press/v119/marx20a.html. ISSN: 2640-3498.

Richard Pace. Six unanswered fair lending questions hindering ai credit model adoption. https://www.paceanalyticsllc.com/post/six-unanswered-fair-lending-questions, June 2022. Updated February 202x; blog post; Pace Analytics Consulting LLC.

Richard Pace. Fool's gold 2: Is there really a low-cost accuracy-fairness trade-off? *The AI LendScape Blog*, Oct 2023.

H Robbins and D Siegmund. On the law of the iterated logarithm for maxima and minima. *Proc. Sixth Berkely Syrup. on Math. Statist. and Probab*, 3:51–70, 1972.

Herbert Robbins. Statistical methods related to the law of the iterated logarithm. *The Annals of Mathematical Statistics*, 41(5):1397–1409, 1970.

Kit T. Rodolfa, Hemank Lamba, and Rayid Ghani. Empirical observation of negligible fairness–accuracy trade-offs in machine learning for public policy. *Nature Machine Intelligence*, 3(10):896–904, October 2021. ISSN 2522-5839. doi: 10.1038/s42256-021-00396-x. URL https://www.nature.com/articles/s42256-021-00396-x. Publisher: Nature Publishing Group.

Cynthia Rudin, Chudi Zhong, Lesia Semenova, Margo Seltzer, Ronald Parr, Jiachang Liu, Srikar Katta, Jon Donnelly, Harry Chen, and Zachery Boner. Amazing things come from having many good models. *arXiv preprint arXiv:2407.04846*, 2024.

Matthew U Scherer, Allan G King, and Marko J Mrkonich. Applying old rules to new tools: Employment discrimination law in the age of algorithms. *SCL Rev.*, 71:449, 2019.

Galen R Shorack. *Probability for statisticians*. Springer, 2000.

DA Sorensen and BW Kennedy. Estimation of response to selection using least-squares and mixed model methodology. *Journal of Animal Science*, 58(5):1097–1106, 1984.

Charles Stein et al. Inadmissibility of the usual estimator for the mean of a multivariate normal distribution. In *Proceedings of the Third Berkeley symposium on mathematical statistics and probability*, volume 1, pp. 197–206, 1956.

Ian Waudby-Smith and Aaditya Ramdas. Estimating means of bounded random variables by betting. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 86(1):1–27, 2024.

Martin Weitzman. *Optimal search for the best alternative*, volume 78. Department of Energy, 1978.

Tijana Zrnic and William Fithian. A Flexible Defense Against the Winner's Curse, November 2024. URL http://arxiv.org/abs/2411.18569. arXiv:2411.18569 [stat].

A LLM Usage.

LLMs were used in the making of this paper as a search and retrieval assistant to generate suggestions for related work or techniques that were useful for proving theorems. They were also used to generate some of the code used in data analys. All LLM suggestions and code were carefully checked for correctness.

B Additional details on empirical analysis.

Dataset preparation. We use pre-selected prediction targets and protected/reference groups given in the datasets. For Folktables, we used data from Alabama from 2018. For HMDA, we use the cleaned dataset given in Cooper et al. (2024) for New York in 2017. Full details of our data cleaning and feature selection are available in our code at: https://anonymous.4open.science/r/lda-6EBA/README.md

The size specifications of our datasets, sub-sampling routines and runs were chosen to produce confident results and demonstrate a plausible approach to implementing the procedure described in this paper.

Datasets versus distributions. An initial challenge of evaluating our method in practice is the absence of ground truth: It is impossible to evaluate predictive models on the distribution from which a dataset was sampled, if we only can access the dataset itself. In light of this difficulty, we treat the dataset itself as representative of a discrete population distribution, and sample iid from this population distribution to arrive at a dataset for training models. First, we define the population distribution to be the empirical measure over the original dataset (Adult, Folktables or HMDA). That is, we define the population distribution $\mathcal{D}$ to be the discrete distribution with equal measure on each of the points in the dataset. Second, we can then generate an iid sample from this distribution by sampling rows from the dataset uniformly at random (i.e., with replacement). Having sampled D this way, we produce a train/test split $D^{\text{train}}, D^{\text{test}}$ of D, where D^{train} is used to define the predictive model and D^{test} is used to evaluate it.Then, for a particular predictive model h trained on D^{train},

Dataset	Logistic Regression	Random Forest	Neural Network
Adult	0.084 (0.013)	0.054 (0.010)	0.095 (0.020)
Folktables	0.056 (0.026)	0.031 (0.017)	0.048 (0.022)
HMDA	0.655 (0.087)	0.416 (0.104)	0.063 (0.088)

Figure 2: Selection rate disparities for each dataset and model class. Reported number is the mean over all runs. Standard deviations are in parentheses.

Dataset	Logistic Regression	Random Forest	Neural Network
Adult	0.824 (0.001)	0.819 (0.003)	0.816 (0.003)
Folktables	0.777 (0.002)	0.808 (0.005)	0.794 (0.004)
HMDA	0.594 (0.003)	0.585 (0.006)	0.505 (0.062)

Figure 3: Accuracy for each dataset and model class. Reported number is the mean over all runs. Standard deviations are in parentheses.

we can compare the estimated disparate impact $\hat{Q}(h, D^{\text{test}})$ (by summing ℓ over D^{test}) against the true population quantity (by summing ℓ over $\mathcal{D}$). In effect, this procedure produces a population distribution so that we can observe $Q(h)$ exactly with respect to $\mathcal{D}$ and characterize the distribution over models. Our algorithm could be run on the full dataset; it would just not allow for comparison with a ground truth sample. When we resampled datasets, we sampled 3000 observations iid.

Just as we cannot observe the population data distribution , we accordingly cannot observe the population model training distribution. That is, we cannot exactly compute probabilities or expectations with respect to $\mathcal{A}(D)$, since our model training process is constituted by a series of possibly complex and opaque operations, and therefore do not lend themselves to closed form computations. Here, we can apply the same strategy as above to generate a population distribution of models: for a given D, we simulate B train/splits and then call the model training procedure on the training data. We then evaluate $\hat{Q}(h_i, D_i^{\text{test}})$ and $Q(h_i)$ as before and save them. This set of values $\{Q(h_i), \hat{Q}(h_i, D_i^{\text{test}})\}_{i=1}^{B}$ can then be used to form a population distribution of model performance and estimated model performance $\mathbb{P}$. When implementing Algorithm 1, we then draw iid samples from $\mathbb{P}$, with replacement.

Model training procedures. We use default parameter settings for each of our model classes, except for the following modifications: For random forests, we fit ten estimators of depth no more than five. For multilayer perceptrons, we fit a model with a single hidden layer of size 25 and run training for a maximum of 600 iterations. The mean and standard deviation of the selection rate disparities and accuracy for each dataset and model class are in Figure 2 and Figure 3, respectively. We train 1000 models for each of 5 different resampled datasets and average the results in all figures. After we generate the population and observed disparate impact for each model, we then resample 1000 times iid from the dataset of model performances. When model miscoverage rates are above the target rate, we suspect it is the result of the relatively small number of models and datasets used to generate the figures.

Dataset	Logistic Regression	Random Forest	Neural Network
Adult	0.000	0.000	0.018
Folktables	0.036	0.099	0.007
HMDA	0.000	0.000	0.135

Figure 4: Miscoverage rates for Algorithm 1 at level 0.05. We compute the number of times our upper bound $\bar{\mu}(\hat{U}_t) \cdot \bar{p}_t(\delta)$ is lower than $g(\hat{U}_t)$ for any t across all runs of the algorithm.

C FORMAL ANALYSIS: DATA-DRIVEN ANYTIME VALID UPPER BOUNDS ON THE CONDITIONAL EXPECTED IMPROVEMENT.

We conclude this section with an analysis of the case in which $\bar{\mu}$ can be estimated from data. In particular, we analyze the case in which the following assumption is satisfied:

Assumption C.1 (Non-decreasing CEI)**.** For all constants $a, a' \in \texttt{supp}\left(\hat{P}_0\right)$ such that $a > a'$ and $a, a' < \texttt{median}(\hat{P}_0)$, it holds with probability 1 that

$$\mathbb{E}_{\hat{P}_0}[a - \hat{Q}_t \mid \hat{Q}_t < a] \geq \mathbb{E}_{\hat{P}_0}[a' - \hat{Q}_t \mid \hat{Q}_t < a']. \tag{7}$$

simultaneously for all $t = 1, 2, \ldots$.

In other words, the expected improvement relative to a on the event that $\hat{Q}_t < a$ is greater than the corresponding quantity with a' for the two thresholds $a > a'$ below the median of P_0.

Note that this assumption cannot automatically be plugged into the $\bar{\mu}$ approach above, since all we know is that the conditional expected improvement is monotonically non-decreasing in a, but we cannot otherwise upper bound it. However, we can infer a high-probability upper bound from the data. We formalize this in Algorithm 2.

At a high level, we leverage the non-decreasing CEI assumption to estimate an upper bound on the CEI that holds with high probability. To do this, we first pick a quantile (one third in our case) of the distribution to estimate, and then we estimate the CEI at a high probability lower bound on that quantile. We then use this as a high-probability choice of $\bar{\mu}$. Not that, unlike above, $\bar{\mu}$ is not an upper bound with probability 1. Thus, we must take a union bound to ensure that overall our guarantee holds with probability at least $1 - \delta$.

The choice of which quantile to threshold on is arbitrary, but it must balance two factors. First, the quantile must not be too close to zero or else we will not have much data to estimate based on, and will thus have wide confidence intervals. Second, the lower the quantile, the smaller the CEI (based on Assumption 3.4), so we will be estimating a smaller quantity for a smaller quantile. Thus, the first factor is about the difference between a high probability lower bound and the true quantity, and the second is about the magnitude of the true quantity. We'd like the combination of these two to be as small as possible. We leave exploration of the optimal choice of the quantile for future work.

Algorithm 2 LDA Search with Adaptive Stopping and CEI Estimation

input: Stopping threshold γ and failure probability δ.

1: Let $T_1 = \lceil 18 \log(3/\delta) \rceil$ and draw T_1 samples $\{\hat{Q}_s\}_{s=1}^{T_1}$.

2: Compute the empirical quantile at level $1/3$:

$$C \triangleq \hat{Q}_{(\lfloor T_1/3 \rfloor)}.$$

3: **for** $t = 1, 2, \ldots$ **do**

4: Draw a new sample $\hat{Q}_t$ and compute $\hat{U}_t = \min_{s \leq t} \hat{Q}_s$.

5: Let $\bar{p}_t$ be defined as in Lemma 3.2. Also, define

$$\begin{aligned} \Delta_t &\triangleq C - \hat{Q}_t \\ S_t &\triangleq \{i \leq t \;:\; \mathbb{I}\{\Delta_i > 0\}\} \\ \bar{\mu}_t &\triangleq \begin{cases} \bar{\mu}_t^{\text{eb}}(\{\Delta_s\}_{s=1}^t, \delta/3, S_t) & \text{if } S_t \neq \varnothing \\ \hat{U}_t & \text{otherwise} \end{cases} \end{aligned}$$

where $\bar{\mu}^{\text{eb}}$ is defined as in Corollary D.6.

6: **if** $\bar{\mu}_t \cdot \bar{p}_t(\delta/3) < \gamma$ **then**

7: **return** $\hat{U}_t$

8: **end if**

9: **end for**

We are now ready to state our theorem.

Theorem C.2. *Under Assumptions 3.4 and C.1, for all $\gamma > 0$ and $\delta > 0$, Algorithm 2 run with $\mathcal{P} = \hat{P}_0$, γ and δ terminates at a time $\tau \in \mathbb{N}$ such that*

$$\mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) \geq 1 - \delta. \tag{8}$$

D DEFERRED PROOFS.

Results are restated before proofs for convenience.

D.1 DEFERRED PROOFS FOR SECTION 3.2

Lemma 3.2. *Let $\{X_t\}_{t=1}^\infty$ be a sequence of iid random variables distributed according to a law $\mathcal{P}_0$. Let $\mathcal{P} \triangleq \mathcal{P}_0^\infty$ be their joint distribution. Let $Y_t \triangleq \min_{s\in[t]} X_s$. For any $\alpha \in (0, 1)$, define*

$$\bar{p}_t(\alpha) = \begin{cases} 1 - e^{-1/\alpha} & \text{if } t = 1 \\ 1 - \left(\frac{(t-1)}{\alpha} + 1\right)^{-1/(t-1)} & \text{otherwise.} \end{cases}$$

Then,

$$\mathcal{P}(\exists t \in \mathbb{N} \ : \ \mathcal{P}_0(X_{t+1} < Y_t \mid Y_t) > \bar{p}_t(\alpha)) \leq \alpha.$$

Proof of Lemma 3.2. At a high level, we proceed as follows:

1. First, show that it suffices to consider the case where the X_t are uniform on $[0, 1]$ via the probability integral transform.
2. Then, we show that it suffices to provide an anytime-valid upper bound on the running minimum of the sequence.
3. Finally, we show that $\bar{p}_t$ as defined above yields such a bound.

We begin by using the probability integral transform to "convert" our X_t's into uniform random variables. Let F_X be the CDF of X_t. Define

$$F_X^{-1}(u) = \inf\{x : F_X(x) \geq u\}.$$

Let $\{U_t\}_{t=1}^\infty$ be iid uniform random variables on $[0, 1]$, defined on $\mathcal{P}$. Let $V_t = \min_{s\in[t]} U_s$. Then, it holds, by e.g. Ch. 6, Theorem 3.1 of Shorack (2000), that

$$\{F_X^{-1}(U_t), F_X^{-1}(V_t)\}_{t=1}^\infty \stackrel{d}{=} \{X_t, Y_t\}_{t=1}^\infty.$$

Because F_X is monotone, F_X^{-1} is monotone as well. Therefore,

$$\begin{aligned}
\mathcal{P}_0[X_{t+1} < Y_t \mid \{X_s\}_{s=1}^t] > \bar{p}_t(\alpha) &= \mathcal{P}_0[X_{t+1} < Y_t \mid Y_t] > \bar{p}_t(\alpha) \\
&= \mathcal{P}_0[F_X^{-1}(U_{t+1}) < F_X^{-1}(V_t) \mid F_X^{-1}(V_t)] > \bar{p}_t(\alpha) \\
&\qquad (\{F_X^{-1}(U_t), F_X^{-1}(V_t)\}_{t=1}^\infty \stackrel{d}{=} \{X_t, Y_t\}_{t=1}^\infty) \\
&= \mathcal{P}_0[F_X^{-1}(U_{t+1}) < F_X^{-1}(V_t) \mid V_t] > \bar{p}_t(\alpha) \\
&\qquad (U_{t+1} \perp \{V_t\};\ F_X^{-1}(V_t) \text{ is measurable with respect to } \sigma(F_X^{-1}(V_t)) \subseteq \sigma(V_t)) \\
&\leq \mathcal{P}_0[U_{t+1} < V_t \mid V_t] > \bar{p}_t(\alpha) \quad (F_X^{-1} \text{ is weakly increasing}) \\
&= V_t > \bar{p}_t(\alpha).
\end{aligned}$$

The last inequality follows from the fact that U_{t+1} is uniformly distributed on $[0, 1]$, so the probability it falls below V_t is precisely V_t. Thus,

$$\mathcal{P}(\exists t \in \mathbb{N} \ : \ \mathcal{P}_0(X_{t+1} < Y_t \mid \{X_s\}_{s=1}^t) > \bar{p}_t(\alpha)) \leq \mathcal{P}(\exists t \in \mathbb{N} \ : \ V_t > \bar{p}_t(\alpha))$$

Thus, our goal is now to provide an anytime-valid upper-bound on V_t.

Define the martingale

$$\begin{aligned}M_t(\theta) &\triangleq \frac{1}{(1-\theta)^t}\mathbb{I}\{V_t \geq \theta\} \\ &= M_{t-1}(\theta) \cdot \left(\frac{1}{1-\theta}\mathbb{I}\{U_t \geq \theta\}\right). \end{aligned} \tag{9}$$

This is a martingale because

$$\begin{aligned}\mathbb{E}[M_t \mid M_{t-1}] &= \mathbb{E}\left[M_{t-1}(\theta)\left(\frac{1}{1-\theta}\mathbb{I}\{U_t \geq \theta\}\right) \,\middle|\, M_{t-1}(\theta)\right] \\ &= \frac{M_{t-1}(\theta)}{1-\theta}\mathbb{E}\left[(\mathbb{I}\{U_t \geq \theta\}) \,\middle|\, M_{t-1}(\theta)\right] \\ &= \frac{M_{t-1}(\theta)}{1-\theta}\Pr[U_t \geq \theta] \\ &= M_{t-1}(\theta).\end{aligned}$$

Moreover, it is a test martingale because it is nonnegative. Next, we use the "method of mixtures" (see, e.g., Robbins, 1970; Waudby-Smith & Ramdas, 2024) to mix M_t with a uniform distribution on θ over $[0,1]$. Intuitively, placing more mass on smaller values of θ gives us sharper bounds for larger values of t. We choose the uniform distribution here for simplicity. In Theorem E.1, we show how to get an asymptotically tight rate.

$$\begin{aligned}M_t^U(\theta) &\triangleq \int_0^1 M_t(\theta)\, d\theta \\ &= \int_0^1 \frac{1}{(1-\theta)^t}\mathbb{I}\{V_t \geq \theta\}\, d\theta \\ &= \int_0^{V_t} \frac{1}{(1-\theta)^t}\, d\theta.\end{aligned} \tag{10}$$

By Fubini's theorem, this is also a test martingale. Applying Ville's inequality (Theorem D.2), for any $\alpha \in (0,1)$,

$$\begin{aligned}\mathcal{P}\left(\exists t \in \mathbb{N} \;:\; M_t^U(\theta) > \frac{1}{\alpha}\right) &\leq \alpha \\ \mathcal{P}\left(\exists t \in \mathbb{N} \;:\; \int_0^{V_t} \frac{1}{(1-\theta)^t} > \frac{1}{\alpha}\right) &\leq \alpha.\end{aligned} \tag{11}$$

Observe that the integrand is nonnegative, so for any sequence $\bar{p}_t$,

$$\int_0^{V_t} \frac{1}{(1-\theta)^t}\, d\theta > \int_0^{\bar{p}_t} \frac{1}{(1-\theta)^t}\, d\theta \iff V_t > \bar{p}_t. \tag{12}$$

Therefore, we can choose $\bar{p}_t(\alpha)$ such that

$$\int_0^{\bar{p}_t(\alpha)} \frac{1}{(1-\theta)^t}\, d\theta = \frac{1}{\alpha}. \tag{13}$$

For $t = 1$,

$$\begin{aligned}\int_0^{\bar{p}_1(\alpha)} \frac{1}{1-\theta}\, d\theta &= \frac{1}{\alpha} \\ -\log(1-\bar{p}_1(\alpha)) &= \frac{1}{\alpha} \\ \bar{p}_1(\alpha) &= 1 - e^{-1/\alpha}.\end{aligned}$$

For $t \geq 2$,

$$\begin{aligned}
\int_0^{\bar{p}_t(\alpha)} \frac{1}{(1-\theta)^t}\, d\theta &= \frac{1}{\alpha} \\
\frac{(1-\bar{p}_t(\alpha))^{t-1}-1}{t-1} &= \frac{1}{\alpha} \\
\bar{p}_t(\alpha) &= 1 - \left(\frac{t-1}{\alpha}+1\right)^{-1/(t-1)}
\end{aligned}$$

Thus,

$$\begin{aligned}
\mathcal{P}\left(\exists t \in \mathbb{N} \;:\; V_t > \bar{p}_t(\alpha)\right) &= \mathcal{P}\left(\exists t \in \mathbb{N} \;:\; \int_0^{V_t} \frac{1}{(1-\theta)^t}\, d\theta > \int_0^{\bar{p}_t} \frac{1}{(1-\theta)^t}\, d\theta\right) \\
&\hspace{20em} \text{(by eq. (12))} \\
&= \mathcal{P}\left(\exists t \in \mathbb{N} \;:\; \int_0^{V_t} \frac{1}{(1-\theta)^t}\, d\theta > \frac{1}{\alpha}\right) \qquad \text{(by eq. (13))} \\
&\leq \alpha, \qquad \text{(by eq. (11))}
\end{aligned}$$

completing the proof. □

We first prove a theorem for general iid random variables bounded in $[0, 1]$.

Theorem D.1. *For all $\gamma, \delta > 0$ and $\mathcal{P}$, Algorithm 1 run with $\mathcal{P}$, γ, δ and any $\bar{\mu}$ satisfying* **??** *as input terminates at a stopping time $\tau \in \mathbb{N}$ such that*

$$\mathcal{P}(\mathbb{E}_{\mathcal{P}}[X_\tau - X_{\tau+1} \mid X_\tau] < \gamma) \geq 1 - \delta.$$

Proof of Theorem D.1. Observe:

$$\begin{aligned}
\mathcal{P}(\mathbb{E}[X_\tau - X_{\tau+1} \mid U_\tau] > \gamma) &= \mathcal{P}(g(X_\tau) > \gamma) \\
&= \mathcal{P}(\mu(X_\tau)p(X_\tau) > \gamma) \\
&\leq \mathcal{P}(\mu(X_\tau)p(X_\tau) > \bar{\mu}(X_\tau)\bar{p}_\tau(\delta)) \\
&\hspace{6em} (\bar{\mu}(X_\tau)\bar{p}_\tau(\delta) \leq \gamma \text{ by the stopping condition}) \\
&\leq \mathcal{P}(\bar{\mu}(X_\tau)p(X_\tau) > \bar{\mu}(X_\tau)\bar{p}_\tau(\delta)) \\
&\hspace{10em} (\mu(X_\tau) \leq \bar{\mu}(X_\tau) \text{ almost surely}) \\
&= \mathcal{P}(p(X_\tau) > \bar{p}_\tau(\delta)) \qquad (\bar{\mu}(u) \geq 0 \text{ for all } u) \\
&\leq \mathcal{P}(\exists t \in \mathbb{N} \;:\; p(X_t) > \bar{p}_t(\delta)) \\
&\leq \delta \qquad \text{(Lemma 3.2)}
\end{aligned}$$

□

Theorem D.2 (Ville's inequality)**.** *Let $M_1, M_2, \ldots$ be a non-negative supermartingale scaled so that $\mathbb{E}M_1 \leq 1$. Then, for any real number α,*

$$P\left(\sup_{t\geq 1} M_t \geq \frac{1}{\alpha}\right) \leq \alpha.$$

D.2 Deferred proofs of Section 3.3

Theorem 3.5. *Under Assumption 3.4, for all $\gamma > 0$ and $\delta > 0$, Algorithm 1 run with $\mathcal{P} = \hat{P}_0$, γ, δ and any $\bar{\mu}$ that satisfies Definition 3.1 for $\hat{P}_0$ terminates at a time $\tau \in \mathbb{N}$ such that*

$$\mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) \geq 1 - \delta. \tag{6}$$

Proof of Theorem 3.5. First, observe:

$$\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] = \mathbb{E}_{\mathbb{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau] + \mathbb{E}_{\mathbb{P}_0}[(U_\tau - \hat{U}_\tau) - (U_{\tau+1} - \hat{U}_{\tau+1}) \mid \hat{U}_\tau].$$

Under Assumption 3.4,

$$\mathbb{E}_{\mathbb{P}_0}[(U_\tau - \hat{U}_\tau) - (U_{\tau+1} - \hat{U}_{\tau+1}) \mid \hat{U}_\tau] \geq 0, \tag{14}$$

for all t with probability 1. Next, observe

$$\mathbb{E}_{\mathbb{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau] = \mathbb{E}_{\hat{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau]. \tag{15}$$

Finally, from Theorem D.1, we have

$$\hat{P}(\mathbb{E}_{\hat{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) \geq 1 - \delta. \tag{16}$$

Putting it all together, we have

$$\begin{aligned}
\mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) &\geq \mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) && \text{(Equation (14))}\\
&= \hat{P}(\mathbb{E}_{\hat{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) && \text{(Equation (15))}\\
&\geq 1 - \delta. && \text{(Equation (16))}
\end{aligned}$$

□

D.3 Deferred proofs for Section 3.4

Theorem C.2. *Under Assumptions 3.4 and C.1, for all $\gamma > 0$ and $\delta > 0$, Algorithm 2 run with $\mathcal{P} = \hat{P}_0$, γ and δ terminates at a time $\tau \in \mathbb{N}$ such that*

$$\mathbb{P}(\mathbb{E}_{\mathbb{P}_0}[U_\tau - U_{\tau+1} \mid \hat{U}_\tau] \leq \gamma) \geq 1 - \delta. \tag{8}$$

Proof of Theorem C.2. Define the following events.

$$\mathcal{E}_0 = \{C \leq \texttt{median}(\hat{P}_0)\}$$
$$\mathcal{E}_1 = \{\mathbb{E}_{\hat{P}_0}[C - \hat{Q}_{\tau+1} \mid C > \hat{Q}_{\tau+1}, C] \leq \bar{\mu}_\tau\}$$

where μ_τ is as defined in algorithm 2.

Notice that, on $\mathcal{E}_0$ and $\mathcal{E}_1$

$$\begin{aligned}
&\mathbb{E}_{\hat{P}_0}[\hat{U}_\tau - \hat{U}_{\tau+1} \mid \hat{U}_\tau > \hat{Q}_{\tau+1}] \leq \mathbb{E}_{\hat{P}_0}[z - \hat{Q}_{\tau+1} \mid z > \hat{Q}_{\tau+1}] && \text{(Assumption C.1)}\\
\implies &\mathbb{E}_{\hat{P}_0}[\hat{U}_\tau - \hat{Q}_{\tau+1} \mid \hat{U}_\tau > \hat{Q}_{\tau+1}] \leq \mathbb{E}_{\hat{P}_0}[C - \hat{Q}_{\tau+1} \mid C > \hat{Q}_{\tau+1}, C]\\
& && (C \in [\hat{U}_\tau, \texttt{median}(P_0)] \text{ a.s. on } \mathcal{E}_0)\\
&\leq \bar{\mu}_\tau. && (\mathcal{E}_1)
\end{aligned}$$

where $\bar{\mu}_\tau$ is defined as in Algorithm 2. Also, define $\mathcal{E}_2 = \{\hat{P}_0(\hat{U}_\tau > \hat{U}_{\tau+1} \mid \hat{U}_\tau) \leq \bar{p}_\tau\}$. Observe that, on $\mathcal{E}_2$,

$$\hat{P}_0(\hat{U}_\tau > \hat{Q}_{\tau+1} \mid \hat{U}_\tau) \leq \bar{p}_\tau(\delta/3).$$

Combining these, we have, by the fact that the algorithm terminated

$$\bar{\mu}_\tau \cdot \bar{p}_\tau(\delta/3) \leq \gamma. \tag{17}$$

By Lemmas 3.2, D.3 and D.4, $\mathcal{E}_0, \mathcal{E}_1$ and $\mathcal{E}_2$ each occur with probability at least $1 - \delta/3$, so by a union bound, their intersection occurs with probability at least $1 - \delta$. □

Lemma D.3. *For all δ, with probability no less than $1 - \delta/3$,*

$$C \leq \texttt{median}(P_0) \tag{18}$$

where C is defined as in Algorithm 2.

Proof of Lemma D.3. Let $\varepsilon = 1/6$ and let $i^* = \lfloor T_1/2 \rfloor$. Note that the event $C \leq \texttt{median}(\hat{P}_0)$ is the same as the event that $i^* \leq \sum_{t=1}^{T_1} \mathbb{I}\left\{\hat{Q}_t \leq \texttt{median}(\hat{P}_0)\right\}$, since this implies that there are at least i^* draws of $\hat{Q}_t$ less than the median. Note that $\mathbb{I}\left\{\hat{Q}_t \leq \texttt{median}(\hat{P}_0)\right\}$ are independent and distributed as Bernoulli random variables with success probability p. Thus,

$$\begin{aligned}
\hat{P}(C > \texttt{median}(\hat{P}_0)) &= \hat{P}\left(i^* > \sum_{t=1}^{T_1} \mathbb{I}\left\{\hat{Q}_t \leq \texttt{median}(\hat{P}_0)\right\}\right) \\
&\leq \exp\left(-\frac{2(i^* - (1/2 - \varepsilon)T_1)^2}{T_1}\right) && \text{(Hoeffding's inequality)} \\
&\leq \exp\left(-2\varepsilon^2 T_1\right) && \text{(Substituting definition of } i^*.\text{)} \\
&\leq \frac{\delta}{3} && \text{(Substituting definition of } \varepsilon \text{ and simplifying.)}
\end{aligned}$$

□

Lemma D.4. *For all δ, with probability at least $1 - \delta/3$, it holds for all $t = 2, 3, \ldots$ simultaneously that*

$$\mathbb{E}_{\hat{P}}[C - \hat{Q}_{t+1} \mid C > \hat{Q}_{t+1}, C] \leq \bar{\mu}_t$$

where $\bar{\mu}_t$ is defined as in Algorithm 2.

Proof of Lemma D.4. We just need to verify that we can apply Corollary D.6. To do this, we need to verify $\Delta \in S_t$ have the same conditional mean.

Define S and $\{i_t\}_t$ analogously to in Corollary D.6:

$$S = \{t \in \mathbb{N} \ : \ \hat{Q}_t < C\}.$$

Define the sequence $S = (t \in \mathbb{N} \ : \ \hat{Q}_t < C)$. To see that all Δ_{i_t} have the same mean conditional on the past, observe,

$$\begin{aligned}
\mathbb{E}_{\hat{P}}[\Delta_{i_t} \mid \Delta_{i_1}, \ldots, \Delta_{i_{t-1}}, C] &= \mathbb{E}_{\hat{P}}[C - \hat{Q}_{i_t} \mid \Delta_{i_1}, \ldots, \Delta_{i_{t-1}}, C] \\
&= C - \mathbb{E}_{\hat{P}}[\hat{Q}_{i_t} \mid \Delta_{i_1}, \ldots, \Delta_{i_{t-1}}, C] \\
&= C - \mathbb{E}_{\hat{P}}[\hat{Q}_{i_t}] && \text{(Independence of } \hat{Q}_{i_t} \text{ conditional on } D\text{)}
\end{aligned}$$

Thus, since $\hat{Q}_{i_t}$ are identically distributed conditional on D, it holds $\mathbb{E}_{\hat{P}_0}\hat{Q}_{i_t} = \mathbb{E}_{\hat{P}_0}\hat{Q}_{i_s}$ for all $s, t \in \mathbb{N}$ so $\{\Delta_{i_t}\}_{t=1}^{\infty}$ have the same mean conditional on the past and C.

Now, on the event that $\hat{Q}_{t+1} < C$, it holds $t + 1 \in S$. Thus, the guarantee holds for $t + 1$. Finally, we plug in $\delta/3$ for α, which yields the desired result:

$$\mathbb{E}_{\hat{P}}[C - \hat{Q}_{t+1} \mid C > \hat{Q}_{t+1}, C] \leq \bar{\mu}_t.$$

□

The following result provides a high probability upper bound for anytime-valid bounded mean estimation.

Theorem D.5 (Theorem 2, Waudby-Smith & Ramdas (2024)). *Suppose there is a constant ν and stochastic process $(X_t)_{t=1}^{\infty} \sim \mathcal{P}$ for some distribution $\mathcal{P}$ with support bounded on $[0, 1]$ such that, for all t,*

$$\mathbb{E}_{\mathcal{P}}(X_t \mid X_1, \ldots, X_{t-1}) = \nu.$$

Let $\mathcal{F}_t = \sigma(\{X_i\}_{i=1}^t)$ be the σ-field induced by $X_1, \ldots, X_t$. Next, consider any sequence $\{\lambda_t\}_{t=1}^{\infty}$ such that for all t, λ_t is $\mathcal{F}_{t-1}$-measurable. Then, for all $\alpha > 0$, with probability at least $1 - \alpha$, it holds for all $t = 1, 2, \ldots$ simultaneously:

$$\nu \leq \frac{\log(2/\alpha) + \sum_{i=1}^t \lambda_i X_i - (X_i - \hat{\nu}_{i-1})^2(\log(1 - \lambda_i) + \lambda_i)}{\sum_{i=1}^t \lambda_i}$$

We state the following corollary Theorem D.5 which states the result for subsequences of random processes (which amounts to a re-indexing) and uses a particular choice of λ_t. This result follows the recommendations for λ_t in Waudby-Smith & Ramdas (2024) and is an empirical Bernstein-type bound.

Corollary D.6. *Suppose there is a constant ν and stochastic process $(X_t)_{t=1}^{\infty} \sim \mathcal{P}$ for some distribution $\mathcal{P}$ with support bounded on $[0,1]$. Define a sequence of subsets S_t such that $S_{t-1} \subseteq S_t$ and $S_t \setminus S_{t-1} \subseteq \{t\}$. Suppose, for all t such that $t \in S_t$ and $i \in S_t$,*

$$\mathbb{E}_{\mathcal{P}}(X_t \mid S_{t-1}) = \nu.$$

For all $\alpha \in (0,1]$, define

$$\lambda_t \triangleq \min\left\{\sqrt{\frac{2\log(2/\alpha)}{\hat{\sigma}_{t-1}^2\,|S_t|\log(1+|S_t|)}}, \frac{1}{2}\right\} \tag{19}$$

where

$$\hat{\nu}_t \triangleq \frac{\frac{1}{2} + \sum_{i\in S_t} X_i}{1+|S_t|}, \text{ and}$$
$$\hat{\sigma}_t^2 \triangleq \frac{\frac{1}{4} + \sum_{i\in S_t}(X_i - \hat{\nu}_i)^2}{1+|S_t|}.$$

Finally, for all t, let

$$\bar{\mu}_t^{eb}(\{X_s\}_{s=1}^t, \alpha, S_t) \triangleq \frac{\log(2/\alpha) + \sum_{i\in S_t} \lambda_i X_i - (X_i - \hat{\nu}_{i-1})^2(\log(1-\lambda_i)+\lambda_i)}{\sum_{i\in S_t}\lambda_i}. \tag{20}$$

Then, with probability at least $1-\alpha$, it holds for all $t = 1, 2, \ldots$ simultaneously:

$$\nu \le \bar{\mu}_t^{eb}(\{X_s\}_{s=1}^t, \alpha, S_t)$$

Proof of Corollary D.6. Define the sequence $S = (t \in \mathbb{N} \ : \ X_t \in S_t)$. Denote by i_t the t-th element of S. Clearly, λ_t is $\mathcal{F}_{t-1}$-measurable. To apply the theorem, we plug in the sequence $\{X_{i_s}\}_{s=1}^{|S_t|}$ as defined in for X_t in Theorem D.5. □

E A sharper upper bound for Lemma 3.2 with an almost matching lower bound.

Theorem E.1. *Let $\{U_t\}_{t=1}^{\infty}$ be a sequence of iid uniform random variables on $[0,1]$. Let $V_t = \min_{s\in[t]} U_s$. For any constant $a > 1$, define*[5]

$$\tilde{p}_t(\delta) \triangleq \min\left\{1, \inf_{q\in[0,e^{-1})}\left\{\int_0^q \frac{1}{(1-\theta)^t}\frac{a-1}{\theta\cdot(\log(1/\theta))^a}\,d\theta \ge \frac{1}{\delta}\right\}\right\}.$$

Then,

$$\Pr\{\exists t\ V_t > \tilde{p}_t(\delta)\} \le \frac{1}{\delta}. \tag{21}$$

Asymptotically,

$$\lim_{t\to\infty}\frac{\tilde{p}_t(\delta)}{\frac{\log\log t}{t}} \in [1, a].$$

Moreover, this is nearly tight: for any sequence $\{q_t\}_{t=1}^{\infty}$,

$$\Pr\{\exists t\ V_t > q_t\} \le \frac{1}{\delta} \implies \lim_{t\to\infty}\frac{q_t}{\frac{\log\log t}{t}} \ge 1.$$

[5] By convention, $\inf \varnothing = \infty$.

Proof. The lower bound follows directly from Robbins & Siegmund (1972, Theorem 1), which states that

$$\Pr\left\{V_t \geq \frac{\log\log t + 2\log\log\log t}{t} \quad \text{i.o.}\right\} = 1.$$

Therefore, for any $\{q_t\}_{t=1}^{\infty}$ such that

$$\lim_{t\to\infty} \frac{q_t}{\frac{\log\log t}{t}} < 1,$$

there is some t^* such that for all $t \geq t^*$, $q_t < \frac{\log\log t + 2\log\log\log t}{t}$. But this means that $V_t > q_t$ infinitely often for $t \geq t^*$, so $\Pr\{\exists t\ V_t > q_t\} = 1$.

For our upper bound, we follow the proof of Lemma 3.2 to define the test margtingale

$$M_t(\theta) \triangleq \frac{1}{(1-\theta)^t}\mathbb{I}\{V_t \geq \theta\}.$$

In Lemma 3.2, we mixed this martingale over the uniform distribution over $[0,1]$ for θ. This lead to an asymptotically loose bound:

$$\begin{aligned}
\bar{p}_t(\delta) &= 1 - \left(\frac{t-1}{\delta} + 1\right)^{-1/(t-1)} \\
&= 1 - \exp\left[-\frac{1}{t-1}\log\left(\frac{t-1}{\delta} + 1\right)\right].
\end{aligned}$$

By Lemma E.2,

$$\begin{aligned}
\bar{p}_t(\delta) &\sim \frac{1}{t-1}\log\left(\frac{t-1}{\delta} + 1\right) \\
&\sim \frac{\log t}{t}.
\end{aligned}$$

To get something asymptotically tight, we need to mix with a distribution that places more mass on very small values of θ. For some constant $a > 1$, consider the distribution

$$\nu(\theta) \triangleq \frac{a-1}{\theta \cdot (\log(1/\theta))^a}$$

defined on $(0, e^{-1})$. This is a valid probability distribution because

$$\begin{aligned}
\int_0^{e^{-1}} \nu(\theta)\, d\theta &= \int_0^{e^{-1}} \frac{a-1}{\theta \cdot (\log(1/\theta))^a}\, d\theta \\
&= (a-1)\int_1^{\infty} u^{-a}\, du && \text{(substitute } u = \log(1/\theta)\text{)} \\
&= (a-1) \cdot \frac{-1}{a-1} u^{-(a-1)}\Big|_1^{\infty} \\
&= u^{-(a-1)}\Big|_{\infty}^{1} \\
&= 1.
\end{aligned}$$

We define our test martingale to be a mixture of M_t over this distribution ν:

$$\begin{aligned}
M_t^N(\theta) &\triangleq \int_0^{e^{-1}} M_t(\theta)\nu(\theta)\, d\theta \\
&= (a-1)\int_0^{\min(e^{-1}, V_t)} \frac{1}{(1-\theta)^t}\frac{1}{\theta(\log(1/\theta))^a}\, d\theta.
\end{aligned}$$

Again, this is a nonnegative martingale by Fubini's theorem, using the fact that $M_t(\theta)$ is a nonnegative martingale as shown in the proof of Lemma 3.2. Applying Ville's inequality (Theorem D.2), for any $\delta \in (0, 1)$,

$$\Pr\left\{\exists t \; M_t^N(\theta) > \frac{1}{\delta}\right\} \leq \delta.$$

Define

$$\tilde{p}_t(\delta) \triangleq \min\left\{1, \inf\left\{q \in [0, e^{-1}) : \int_0^q \frac{1}{(1-\theta)^t} \frac{a-1}{\theta \cdot (\log(1/\theta))^a} \, d\theta \geq \frac{1}{\delta}\right\}\right\}.$$

For sufficiently large t, the set over which we are taking the infimum will be nonempty, and for such t,

$$\int_0^{\tilde{p}_t(\delta)} \frac{1}{(1-\theta)^t} \frac{a-1}{\theta \cdot (\log(1/\theta))^a} \, d\theta = \frac{1}{\delta}$$

By a simple monotonicity argument,

$$\begin{aligned}
\{V_t > \tilde{p}_t(\delta)\} &\iff \{V_t > \tilde{p}_t(\delta), \tilde{p}_t(\delta) < e^{-1}\} \\
&\implies (a-1) \int_0^{\min(e^{-1}, V_t)} \frac{1}{(1-\theta)^t} \frac{1}{\theta(\log(1/\theta))^a} \, d\theta > \frac{1}{\delta} \\
&\iff \left\{M_t^N > \frac{1}{\delta}\right\}
\end{aligned}$$

Therefore,

$$\Pr\left\{\exists t \; V_t > \tilde{p}_t(\delta)\right\} \leq \delta,$$

which proves eq. (21).

Finally, to prove the asymptotic bounds,

$$\begin{aligned}
g(v, t) &\triangleq e^{-v} \frac{1}{(1-v/t)^t} \frac{1}{(1-\frac{\log v}{\log t})^a} \\
g_1(v, t) &\triangleq e^{-v} \frac{1}{(1-v/t)^t} \\
g_2(v, t) &\triangleq \frac{1}{(1-\frac{\log v}{\log t})^a}.
\end{aligned}$$

By definition, $g = g_1 g_2$. Consider our integral

$$\int_0^{\tilde{p}_t(\theta)} \frac{1}{(1-\theta)^t} \frac{a-1}{\theta \cdot (\log(1/\theta))^a} \, d\theta$$

Make the substitution $v = t\theta$. Then, this becomes

$$\int_0^{t\tilde{p}_t(\theta)} \frac{1}{(1-v/t)^t} \frac{a-1}{v/t \cdot (\log(t/v))^a} \frac{dv}{t} = \frac{a-1}{(\log t)^a} \int_0^{t\tilde{p}_t(\theta)} \frac{1}{(1-v/t)^t} \frac{1}{v \cdot (\log(t/v))^a} \, dv$$

Intuitively, our goal will be to show that this integrand is approximately e^v/v. To do so, observe that for $v > 1$,

$$\begin{aligned}
g_1(v, t) &= \frac{e^{-v}}{(1-v/t)^t} \\
&= e^{-v-t\log(1-v/t)} \\
&\geq e^{-v+t(v/t)} && (-\log(1-x) \geq x \text{ for } x > 0) \\
&= 1 \\
g_2(v, t) &= \frac{1}{\left(1-\frac{\log v}{\log t}\right)^a} \\
&\geq 1 \\
g(v, t) &= g_1(v, t) g_2(v, t) \\
&\geq 1.
\end{aligned}$$

We can now break up our original integral into two parts:

$$\begin{aligned}\frac{a-1}{(\log t)^a}\int_0^{t\tilde{p}_t(\delta)} \frac{1}{(1-v/t)^t}\frac{1}{v(1-\frac{\log v}{\log t})^a}\,dv &= \frac{a-1}{(\log t)^a}\int_0^{t\tilde{p}_t(\delta)} \frac{e^v}{v} g(v,t)\,dv \\ &= \frac{a-1}{(\log t)^a}\int_0^{1} \frac{e^v}{v} g(v,t)\,dv \\ &+ \frac{a-1}{(\log t)^a}\int_1^{t\tilde{p}_t(\delta)} \frac{e^v}{v} g(v,t)\,dv \end{aligned} \tag{22}$$

We'll show that the first of these terms approaches 0. This is because for $v \leq 1$,

$$\begin{aligned} g(v,t) &= e^{-v}\frac{1}{(1-v/t)^t}\frac{1}{(1-\frac{\log v}{\log t})^a} \\ &\leq \frac{1}{(1-1/t)^t}\frac{1}{(1-\frac{\log v}{\log t})^a} \\ &\leq 4 && (t \geq 2; \log v < 0)\end{aligned}$$

Moreover, $g(v,t) \geq 0$. Therefore,

$$\frac{a-1}{(\log t)^a}\int_0^1 \frac{e^v}{v} g(v,t)\,dv \leq \frac{4(a-1)}{(\log t)^a},$$

which goes to 0 as $t \to \infty$. This means that asymptotically,

$$\frac{a-1}{(\log t)^a}\int_1^{t\tilde{p}_t(\delta)} \frac{e^v}{v} g(v,t)\,dv = \frac{1}{\delta} - o(1). \tag{23}$$

Using the fact that $g(v,t) \geq 1$ for $v > 1$,

$$\frac{1}{\delta} \geq \frac{a-1}{(\log t)^a}\int_1^{t\tilde{p}_t(\delta)} \frac{e^v}{v} g(v,t)\,dv \geq \frac{a-1}{(\log t)^a}\int_1^{t\tilde{p}_t(\delta)} \frac{e^v}{v}\,dv.$$

Consider the sequence z_t implicitly defined (for sufficiently large t) as

$$\frac{a-1}{(\log t)^a}\int_1^{z_t} \frac{e^v}{v}\,dv = \frac{1}{\delta}.$$

Clearly, $z_t \geq \tilde{p}_t(\delta)$ because the integrand e^v/v is nonnegative. We will show that $z_t \sim a \log\log t/t$.

The exponential integral Ei is defined

$$\mathrm{Ei}(x) \triangleq \int_{-\infty}^{x} \frac{e^v}{v}\,dv.$$

Therefore,

$$\frac{a-1}{(\log t)^a}\int_1^{z_t} \frac{e^u}{u}\,du = \frac{a-1}{(\log t)^a}(\mathrm{Ei}(z_t) - \mathrm{Ei}(1)).$$

By definition of z_t, for all t,

$$\frac{\delta(a-1)}{(\log t)^a}(\mathrm{Ei}(z_t) - \mathrm{Ei}(1)) = 1.$$

Therefore,

$$\begin{aligned}\lim_{t\to\infty}\frac{\delta(a-1)}{(\log t)^a}(\mathrm{Ei}(z_t) - \mathrm{Ei}(1)) &= 1 \\ \lim_{t\to\infty}\frac{\delta(a-1)}{(\log t)^a}\,\mathrm{Ei}(z_t) &= 1\end{aligned}$$

We can write this with the asymptotic relation

$$\delta(a-1)\,\mathrm{Ei}(z_t) \sim (\log t)^a.$$

By the lower bound shown above, we must have $z_t \geq \tilde{p}_t(\delta) = \Omega(\log\log t/t)$, meaning $z_t \to \infty$. By Lemma E.3, if $z_t \to \infty$, then $\mathrm{Ei}(z_t) \sim e^{z_t}/z_t$. Therefore,

$$\begin{aligned}
\delta(a-1)\,\mathrm{Ei}(z_t) &\sim (\log t)^a \\
\delta(a-1)\frac{e^{z_t}}{z_t} &\sim (\log t)^a \\
e^{z_t - \log z_t} &\sim \frac{(\log t)^a}{\delta(a-1)} && \text{(Lemma E.3, since } z_t \to \infty\text{)} \\
z_t - \log z_t &\sim \log\left(\frac{(\log t)^a}{\delta(a-1)}\right) && \text{(both sides go to } \infty\text{)} \\
z_t &\sim a \log\log t
\end{aligned}$$

Because $\tilde{p}_t(\delta) \leq z_t$,

$$\lim_{t\to\infty} \frac{\tilde{p}_t(\delta)}{\frac{\log\log t}{t}} \leq a.$$

The lower bound we began with yields

$$\lim_{t\to\infty} \frac{\tilde{p}_t(\delta)}{\frac{\log\log t}{t}} \geq 1,$$

completing the proof. □

Lemma E.2. *For a sequence $\{a_t\}_{t=1}^{\infty}$, if $\lim_{t\to\infty} a_t = 0$, then*

$$1 - e^{a_t} \sim -a_t.$$

Proof. We must show that

$$\lim_{t\to\infty} \frac{1-e^{a_t}}{-a_t} = 1. \tag{24}$$

We proceed as follows.

$$\begin{aligned}
\lim_{t\to\infty} \frac{1-e^{a_t}}{-a_t} &= \lim_{t\to\infty} \frac{e^{a_t}-1}{a_t} \\
&= \lim_{u\to 0} \frac{e^u - 1}{u} && (\lim_{t\to\infty} a_t = 0) \\
&= \lim_{u\to 0} \frac{e^{0+u} - e^0}{u} \\
&= \frac{d}{du} e^u \bigg|_{u=0} = 1.
\end{aligned}$$

□

Lemma E.3. *As $z \to \infty$,*

$$\mathrm{Ei}(z) \sim \frac{e^z}{z}.$$

Proof.

$$\begin{aligned}
\lim_{z\to\infty} \frac{\mathrm{Ei}(z)}{\frac{e^z}{z}} &= \lim_{z\to\infty} \frac{\frac{d}{dz}\mathrm{Ei}(z)}{\frac{d}{dz}\frac{e^z}{z}} \\
&= \lim_{z\to\infty} \frac{\frac{e^z}{z}}{\frac{ze^z - e^z}{z^2}} \\
&= \lim_{z\to\infty} \frac{1}{\frac{z-1}{z}} \\
&= 1.
\end{aligned}$$

□