Transportation Research Part C 151 (2023) 104118

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Behavioral inference from non-stationary policies: Theory and application to ridehailing drivers during COVID-19 lockdowns

Matthew Battifarano [a], Sean Qian [a,b,*]

[a] Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, 15213, PA, USA
[b] Heinz College, Carnegie Mellon University, Pittsburgh, 15213, PA, USA

ARTICLE INFO

Keywords:
Sequential hypothesis testing
E-process
Transportation network companies
COVID-19

ABSTRACT

In the aftermath of a disruptive event like the onset of the COVID-19 pandemic, it is important for policymakers to quickly understand how people are changing their behavior and their goals in response to the event. Choice modeling is often applied to infer the relationship between preference and behavior, but it assumes that the underlying relationship is stationary: that decisions are drawn from the same model over time. However, when observed decisions outcomes are non-stationary in time because, for example, the agent is changing their behavioral policy over time, existing methods fail to recognize the intent behind these changes. To this end, we introduce a non-parametric sequentially-valid online statistical hypothesis test to identify entities in the urban environment that ride-sourcing drivers increasingly sought out or avoided over the initial months of the COVID-19 pandemic. We recover concrete and intuitive behavioral patterns across drivers to demonstrate that this procedure can be used to detect behavioral trends as they are emerging.

1. Introduction

When we observe a sequence of binary decision *outcomes* generated by agents involved in *active learning* (Cohn et al., 1996) can we deduce the reward signal they appear to be chasing? Such problems arise naturally when we are able to observe agent behavior but not the intent that generated the behavior. When the *policy* the agent is using to select their actions does not change over time, existing methods such as inverse reinforcement learning (Arora and Doshi, 2021) and discrete choice modeling (Greene, 2009) are available to identify a reward function which best rationalizes the observed behavior. Existing methods, to the extent that they consider uncertainty, consider agents who have only *aleatory uncertainty* (Smith, 2013, Definition 1.6) about the relationship between action, environment, and reward. As a result, these methods assume that successive observations of state, action, and reward carry no information about the environment that is not already known to the agent and will not affect their view of the world. In other words, when we use existing methods we assume that the agent whose behavior we are observing is no longer learning anything about their environment. However, it is often unreasonable to assume that an agent has no *epistemic uncertainty* (Smith, 2013, Definition 1.7) about the relationship between state, action, and reward; particularly if the agent is interacting with a complex environment with a high-dimensional action space. In these cases, the observation of an outcome (and corresponding reward) will influence the agent's policy moving forward; in other words, the agent is learning successively better policies over time and we cannot apply existing methods to rationalize their behavior. Given a sequence of *outcomes* generated by an agent, we develop a sequentially-valid test statistic which reports when in the observation window, if ever, there is enough evidence to reject the notion that the agent is

* Corresponding author at: Department of Civil and Environmental Engineering, Carnegie Mellon University, Pittsburgh, 15213, PA, USA.
E-mail addresses: mbattifa@andrew.cmu.edu (M. Battifarano), seanqian@cmu.edu (S. Qian).

https://doi.org/10.1016/j.trc.2023.104118
Received 20 July 2022; Received in revised form 28 January 2023; Accepted 28 March 2023
Available online 12 April 2023

not seeking (avoiding) the outcome at a pre-specified confidence level. Notably, the test statistic does not require any knowledge of the action space. We demonstrate the test statistic on simulated interactions with a multi-armed bandit before applying it to the behavior of real ride-sourcing drivers before and during the COVID-19 lockdown in the city of Pittsburgh, Pennsylvania.

This methodology is most useful when we wish to rationalize observed decision making when it is likely that the decision-makers themselves are modifying the criteria they use to make decisions *in response to* the outcomes of their decisions over time. In other words, the agents refine their policies to reduce their *epistemic uncertainty* about the relationship between state, action, and reward. Conventional notions of choice modeling assume a time-independent criteria and thus will mistake the reduction in epistemic uncertainty for a large degree of aleatory uncertainty. In contrast, our methodology aims to first identify *if* the agents appear to be meaningfully changing their strategy over the observation period and if so, to then identify the region of the decision outcome space the agents appear to be attracted to over the observation period.

In particular, in this work we are interested in making a *statistical claim* about the how ride-sourcing driver were changing their behavior based on scant information about their decisions, and next to no information about the criteria they used to make them, in an environment whose reward structure was highly variable and was highly uncertain to the drivers themselves. We offer an alternative approach to choice modeling or inverse reinforcement learning in this setting that works in real-time on relatively few assumptions but asks a fundamentally different question: what is changing about their behavior? and, how soon can we detect a change with high probability?

The methodology we develop here may be directly applied to a more general class of problems, to answer the question: is an event becoming more (less) likely over time? In this paper we are interested in detecting changes in behavior so it is useful to conceptualize the event as being generated by a learning agent interacting with an environment, but our process makes *no* assumptions on the data-generating process. Our only requirements are that the outcomes are observed sequentially in chronological order and that past outcomes do not depend on future outcomes. In this sense our method belongs to a wider class of *online changepoint detection* methods capable of identifying the moment when there is enough evidence to suggest that the probability of an event may be non-stationary in time or, if more specificity is required, that this probability may be increasing (decreasing) in time.

Importantly, this approach is specifically designed to be applied to circumstances in which the researcher is not willing to assume that a decision making process is stationary in time. In particular, when the observed decision maker is changing their decision making process in response to the outcome of previous decisions. More broadly, as we discuss, this method may be applied to detect when the frequency of some event becomes non-stationary. Ours is not a replacement for discrete choice modeling, rather an additional tool to leverage when the stationarity assumption cannot be met. Explicitly, this approach asks a fundamentally different question of a sequence of observed decisions than choice models do; it does not estimate a utility function nor a policy for the agent, rather it identifies *changes* in observed behavior which may indicate what aspects of their environment the agent is learning to avoid or learning to seek out.

The remainder of this paper is organized as follows. Section 2 discusses the prior work on which this work relies. In particular, we review choice modeling with emphasis on applications to transportation research and draw connections to inverse reinforcement learning. We then discuss prior work in sequential statistical hypothesis testing which our work directly leverages. Although sequential hypothesis testing has been known in statistics for many years it is currently an active area of statistics research. Section 3 introduces our methodologies. We first detail our original methodology to extract a sequence of stops from OpenStreetMaps feature tag data and the ride-sourcing driver GPS trajectories. We then introduce our test statistic, the main contribution of this paper. Section 4 introduces and describes the datasets we use in this paper. In Section 5 we then offer a descriptive analysis of the ride-souring GPS trajectory data with a particular emphasis on comparisons between 2019 and 2020 trajectory data. In Section 6 we describe the set-up for our experiments. We first demonstrate the test statistic on two simulated datasets before describing how we apply the test statistic to our ride-souring GPS dataset. Section 7 reports the main results from the sequential hypothesis test on the ride-sourcing driver data. It is followed in Section 8 by a detailed quantitative and qualitative analysis of the results and their implications. In Section 9 we discuss some immediate applications of our work here and some potential future applications of the methodology. Finally in Section 10 we offer concluding remarks and directions of future work.

2. Background and prior work

2.1. Choice modeling in transportation systems

Inferring intent from behavior is an important theme in transportation research. Applications of discrete choice modeling to transportation systems first appeared in the 1960's and were aimed primarily at questions of mode choice, expanding over the course of the 1970's to other dimensions of travel choice, including car ownership, travel demand, and housing (Ben-Akiva, 1985; McFadden, 1974). One classical methodology in this line of inquiry is the multinomial logit model (MNL) (McFadden et al., 1973; Bhat, 1997) which assigns to each choice in an exhaustive set of mutually exclusive choices a probability:

$$\Pr[Y = j] \propto \exp\left(\langle X_j, \beta \rangle\right) \tag{1}$$

where Y is a random variable representing the decision reached, X_j represents a vector of attributes of the decision j, and β a vector of *preference parameters*. The dot product between the preference parameters β and the characteristics of the decision X_j is known as the *utility* or *reward* of the choice j. It is assumed that agents wish to take the decision which maximizes their utility. The particular probabilistic model defined by Eq. (1) arises from an assumption that the observed utility, $U_j := \langle X_j, \beta \rangle$, differs from the

true value of utility, V_j, (used by the agent to make the decision) by a Gumbel-distributed random quantity, assumed to be i.i.d. over both the choice set and decisions.

Typically, random utility theory is motivated from the perspective of the statistician: it is assumed the agent has access to the full set of decision features and is able to correctly select the highest utility choice. The randomness that we, the statisticians, observe is due entirely to the fact that we cannot observe the full set of criteria the agents are using. However, the randomness can also be thought of as arising from the environment which motivates random utility theory from the agent's perspective. In this view, the true utility, V_j, is revealed only after the agent has made their choice, and the observed utility, U_j, represents the agent's *anticipated* utility from taking decision j. In order to arrive at Eq. (1), the difference between what the agent anticipated and what they actually received is Gumbel distributed, assumed to be i.i.d. over both the choice set and decisions.

When viewed from the agent's perspective, it is clear that the agent is assumed to know quite a lot about how the environment will respond to their action. In fact, the agent has a "correct" model of the environment in that their uncertainty about the reward is entirely *aleatory*, stemming from inherent randomness rather than from misspecification. When the agent is interacting with a complex environment with a large action space, it may not be reasonable to assume that the agent has no epistemic uncertainty about their environment. In such cases, the agent takes actions based on their *beliefs* about the reward which are either substantiated or refuted by the actual reward they receive and may lead the agent to adjust their beliefs for subsequent decisions. As a result there is a temporal dependence between actions which standard choice modeling cannot account for.

In typical transportation applications of choice modeling, the dataset of decisions is aggregated across individuals at a point in time; any temporal dependence that may exist within an individual's policy over time is not detrimental to the application. However, as disaggregate data, for example individual GPS trajectories, becomes both more prevalent and more detailed, the application of choice modeling to datasets of individual's behavior over time becomes more enticing. *Dynamic discrete choice models* (DDC) (see Keane and Wolpin, 2009) are one effective way to model sequential decisions; *inverse reinforcement learning* (IRL) is another.

2.2. Dynamic discrete choice models and inverse reinforcement learning

Dynamic discrete choice models and inverse reinforcement learning both observe a *trajectory* of behavior and attempt to estimate produces a utility or reward function which best rationalizes the observed behavior. The dataset of behavior in this context are *sequences* of state, action pairs for each observed individual. Maximum Entropy IRL (MaxEnt IRL) (Ziebart et al., 2008) effectively extends Eq. (1) to the Markov decision process (MDP) on which reinforcement learning is based:

$$\Pr[\zeta = (s_1, a_1, \dots, s_T, a_T)] \propto \exp\left(\sum_{t=1}^{T} r_\theta(s_t, a_t)\right) \tag{2}$$

where ζ is a random variable representing a trajectory through the MDP, r_θ is the reward (parameterized by θ) as a function of the state at time t and the action taken at time t.

In both DDC and IRL, the decision-makers are forward-looking in the sense that the value of a decision at time t influences the value of future decisions. When the underlying process is assumed to be an MDP, the action taken at each time step will influence the state the agent finds themselves in the next time step which will in turn, affect the value of the actions at the next time step, and so on. In principle, this approach is general enough to accommodate an aspect of "learning", because the notion of state may be expanded to additionally represent the entire past trajectory of decisions. The reward or utility would then a function of the agent's experience in addition to the state of the agent and the action they choose, effectively replacing $r_\theta(s_t, a_t)$ with $r_\theta(s_t, a_t, s_{t-1}, a_{t-1}, \dots, s_0, a_0)$. Such a modeling choice would, in principle, enable the practitioner to estimate both the actual reward or utility function (i.e. the reward or utility the agent *actually* receives for taking an action in a state) and the flawed estimate of the reward function they were actually using to make decisions based on their past experience. However, there is a theoretical problem with this approach caused by the fact that the trajectory contains a lot of information about the current state. In fact, in the case where the underlying process is an MDP, the distribution from which s_t was drawn is completely determined by just s_{t-1} and a_{t-1}. As a result of the dependence between current state and the experience, it will be difficult for the fitted policy to correctly attribute an action to the current state (i.e. the true reward) or to experience (i.e. the agent's estimated reward). This can be understood in analogy to multicollinearity in linear regression. Broadly speaking, this problem is a kind of identifiability problem because a state-conditioned action that is infrequently observed may have been observed infrequently because its expected state-conditioned reward is truly small and the agent know this fact or because its expected state-conditioned reward is truly large and the agent has yet to learn this fact. Both conclusions are supported equally by the data and, without additional information or assumptions, it is not possible to identify which one better represents the decision making process under study.

In all cases the reward function (and by extension, the policy) is explicitly assumed to be time independent, as the parameters of the reward function are fixed over time. In other words, the agent's beliefs about the relationship between state, action, and reward are unchanged by their previous observations of rewards. The assumption of a time-independent policy in IRL is typically motivated from the perspective of *behavioral cloning* (Sutton et al., 1998) wherein the behavior is assumed to be an "expert demonstration" of the task. The experts, it is assumed, have already learned everything they need to know about the task. In this sense, inverse reinforcement learning as a term used to describe the process of inferring a reward function from demonstrations of a fixed expert policy can be considered somewhat of a misnomer: if (forward) reinforcement learning is the process of efficiently balancing exploration and exploitation to learn a policy that maximizes a known reward function, then inverse reinforcement learning should refer to the process of inferring the reward function from the behavior of agents who are in the process of efficiently learning policies to maximize it.

Perhaps the best way to understand why choice modeling struggles when decision-makers are learning is by the following analogy. Suppose you are trying to learn the objective of a game you know nothing about solely by watching the moves of a novice player (over one or more rounds) without the ability to ask them questions or even know if they won or lost. If the player were an "expert" it would be possible to estimate a reward function that induces a policy that maximizes the likelihood of the observed decisions. This is the domain of choice modeling. However, the player is a novice who is also learning (and hopefully getting better) as you watch them play. Learning a policy (either directly or indirectly via a learned reward/utility function) that reproduces the observations is of limited use because we know, generally speaking, moves from earlier games are worse than moves from later games, but we do not know how much worse because we were not able to observe the reward. In particular, we know that each move (or round) causes the player to revise their policy based on the reward, but not in a way that we can estimate from observations of the moves in isolation because we do not have access to the reward signal that the player is using to learn. Our approach side-steps these problems in this setting by effectively identifying moves that are being played less often over time (which we might assume are bad moves) and identifying moves that are being played more often over time (which we might assume are good moves). While our approach yields neither an estimated policy nor an estimated utility function at all, it does yield information about how the novice is changing their strategy as they gain more experience.

2.2.1. Choice modeling and transportation network companies

Choice models have an important role to play in understanding the impact of TNCs on transportation systems. On the demand side, where much of the research has been performed, choice modeling helps transportation planners understand what motivates travelers to use ride-sourcing over other modes of transportation. One of the first demand-side choice models specifically aimed at ride-sourcing systems demand was given by Dias et al. (2017) who estimated a probit choice model to explain the influence of socio-economic and demographic factors on the frequency of ride-sourcing use. Alemi et al. (2018) fit a binary logit model to estimate the extent to which socio-demographic, attitudinal, and environmental factors influence traveler adoption of ride-hailing services. Similarly, Lavieri and Bhat (2019) construct a latent-variable model relating socio-demographic and revealed transportation preferences to the adoption and utilization of ride-hailing services.

Choice models have also been applied to study ride-sourcing drivers. Hall and Krueger (2018) provide the first comprehensive study of the labor market for Uber drivers. Though their findings offer a descriptive account of Uber drivers, Hall and Krueger (2018) identify important features that distinguish Uber drivers from the general population. Their findings lay the foundation for subsequent research aimed at explaining driver behavior. Berliner and Tal (2018) estimate an ordinal logit model to understand how socio-demographic and attitudinal factors influence willingness to drive for a TNC. de Ruijter et al. (2022) model the emergent behavior of decentralized driver decision making on the evolution of supply in ride-hailing networks. Bansal et al. (2020) estimate preferences of both TNC riders and drivers to understand both TNC usage and vehicle purchasing behavior.

There is also a small but growing body of research in applying choice models to understand TNCs in the context of urban delivery. Miller et al. (2017) model traveler willingness to work as a delivery driver on a TNC platform. Punel and Stathopoulos (2017) model the preferences of package senders engaged with TNC goods delivery platforms.

2.2.2. Research contributions

Our contribution in this area is to extend choice modeling to apply in circumstances where agents engaged in sequential decision making are likely changing their decision-making criteria over the course of the observation period in pursuit of maximizing some reward or utility function. We introduce a novel statistical framework to detect whether (and when) learning may be occurring and, if it is, identify regions of the decision-criteria space that the agents are drawing towards over time.

2.3. Inferring activity from GPS trajectories

In this paper we introduce an original stop inference procedure to infer a sequence of what we call "service stops" from driver GPS trajectories. This procedure is an example of the more general process of inferring activity from GPS trajectories which is an active area of research. We note that our procedure, though an original contribution of this paper, is not the main contribution, nor intended to be a state-of-the-art activity inference procedure. Nevertheless, we provide a brief survey of recent work in this area.

Servizi et al. (2020) combine GPS trajectories and OpenStreetMaps data to train neural networks to detect stops within smartphone GPS trajectories using geospatial context from OpenStreetMaps. Kim et al. (2022) tackle a broader inference problem, learning the travel mode from GPS trajectories via recurrent convolutional neural networks. Xiao et al. (2016) develop a neural network to tackle the related problem of inferring the purpose of a trip from the GPS traces.

Our methodology attempts to segment low-speed GPS pings into stops via the application of a simple probabilistic model. In contrast to more general travel mode inference or trip purpose inference, our methodology is concerned exclusively with detecting a stop, independent of its purpose. We prioritize the simplicity of the model and the interpretability of the hyperparameters in our approach. Importantly, in contrast to Servizi et al. (2020), we rely only on the GPS trajectory itself to perform the inference. It is likely useful to consider contextual information in stop inference but because our statistical procedure will measure the frequency with which stops are nearby certain OpenStreetMaps entities, it was important for us to avoid introducing a artificial dependence between OpenStreetMaps data and the stops we infer.

2.4. Sequential hypothesis testing with non-negative super-martingales

Sequential hypothesis testing extends the notion of classical hypothesis testing to the setting in which data may be observed, and the hypothesis test may be performed, in sequence. The value of sequential hypothesis testing methods is that, unlike conventional hypothesis testing methods, they are valid under arbitrary data-dependent stopping times (Ramdas et al., 2020). That is, the researcher can inspect the data and results of the hypothesis test over time and decide when to stop collecting data based on any function of the data collected so far, including the value of the test statistic itself. When using conventional hypothesis testing, using the data to determine when to stop the test and draw conclusions invalidates the results and is known as "p-hacking" (Head et al., 2015). The pursuit of sequential hypothesis testing and analogous confidence *sequences* is the pursuit of "anytime-valid" tests: tests that retain their validity in the face of arbitrary data-dependent stopping times.

An intuition for sequential hypothesis testing can be developed from gambling. In fact some of the earliest academic work in probability theory developed in order to define what it meant for a gamble to be "fair" (Shafer and Vovk, 2019, Preface). This legacy of gambling remains when we use a "fair coin" or "fair dice" to exhibit some aspect of probability theory, but as Shafer and Vovk (2019) expound fairness is not an axiom, but rather a result of the game being played. Fairness in gambling is fundamentally a question of pricing. If a gambler is being offered a ticket that may be redeemed for \$1 if the outcome of the next coin toss is heads, what is the most the gambler should be willing to pay for it? The "fair" price in this instance is exactly the probability that the coin toss results in heads. If the price of a ticket is more than the "fair" amount, we would think the gambler is being cheated; if the price is less than the "fair" amount, we would think the house is being cheated.

Shafer and Vovk (2019) refine this idea and give it a rigorous mathematical foundation by developing a game-theoretic theory of probability. However, what Shafer and Vovk develop is also true in the more conventional measure-theoretic theory of probability. We suppose that a gambler will play this coin toss game repeatedly, beginning with \$1 and without the ability to borrow money or go into debt at any point. We can represent the gambler's wealth over the rounds of the game (time) as a sequence of random variables (a stochastic process). If the gambler is allowed to purchase or short-sell the ticket for heads in any quantity then Shafer and Vovk (2019, Proposition 1.2) prove that the gambler has a strategy over repeated plays of this game which guarantees that either the price of the game is fair (i.e. the price matches the probability) or the gambler's wealth will grow infinitely.

When the coin toss game is fair, the gambler's wealth over time is a non-negative (super-) martingale (NSM). A martingale is a sequence of random variables whose value at any time is the conditional expectation of its value at any future time (Ramdas et al., 2020). A non-negative martingale is a martingale for which every random variable in the sequence is non-negative with probability 1. For example, when the coin toss game is fair, the expected value of the gambler's wealth after the next coin toss given their wealth now is simply the gambler's wealth now. In other words, the expected value of the gambler's winnings on each round is equal to the price they paid for the ticket: it is a fair game. In the language of mathematical probability, let $\mathcal{F}_t$ be a filtration and P be a measure on $\mathcal{F}_\infty$. Following Ramdas et al. (2020), a stochastic process $(M_t)_{t=0\ldots}$ is a **martingale** if M_t is $\mathcal{F}_t$-measurable, P-integrable, and

$$\mathbb{E}_P[M_t \mid \mathcal{F}_s] = M_s \quad \text{for any } t \text{ and } s \le t. \tag{3}$$

$(M_t)_{t=0,\ldots}$ is a **super-martingale** when instead,

$$\mathbb{E}_P[M_t \mid \mathcal{F}_s] \le M_s \quad \text{for any } t \text{ and } s \le t. \tag{4}$$

Finally, $(M_t)_{t=0,\ldots}$ is a **sub-martingale** when instead,

$$\mathbb{E}_P[M_t \mid \mathcal{F}_s] \ge M_s \quad \text{for any } t \text{ and } s \le t. \tag{5}$$

The filtration $\mathcal{F}_t$ can be described in conceptual terms as the set of all information available at time t. The *canonical filtration* for a stochastic process M_t is usually taken to be the σ-algebra generated by the sequence $(M_i)_{i=1}^t$, which is to say the set of all possible realizations of the sequence $(M_i)_{i=1}^t$. Although the filtration is an important concept to the theory of stochastic processes, in practice we typically do not have to think much about it so long as we are careful that the information available at time t does not include information from future times. In this paper we will abbreviate the conditional expectation $\mathbb{E}_P[M_t \mid \mathcal{F}_s]$ as $\mathbb{E}_s[M_t]$ when both the probability measure P and the filtration $\mathcal{F}_s$ are clear from context.

Non-negative super-martingales are at the heart of several converging lines of statistical research in sequential hypothesis testing and sequential estimation (i.e. confidence sequences) (Howard and Ramdas, 2019). Shafer et al. (2011) refer to "test martingales", Shafer and Vovk (2019), Shafer (2019) to "betting scores", and Vovk and Wang (2021), Grünwald et al. (2020) to "safe e-values". Underlying all of this work are non-negative super-martingales. In fact, Ramdas et al. (2020) show that *every* sequential hypothesis test or estimation method must either itself rely on non-negative super-martingales or be strictly out-performed by one.

2.4.1. Ville's inequality

The workhorse of hypothesis testing with non-negative super-martingales is *Ville's Inequality* (Ville, 1939) (see also Howard et al., 2020, Section 2.3) which extends Markov's inequality to stochastic processes, namely non-negative super-martingales. Let Y_t be a non-negative super-martingale adapted to the filtration $\mathcal{F}_t$; Ville's inequality allows us to conclude that

$$\Pr[\exists t : Y_t \ge \alpha \mathbb{E}[Y_0]] \le \frac{1}{\alpha} \tag{6}$$

where Y_0 is the initial value of the NSM. What Ville's inequality states is that the event that a NSM exceeds its initial (expected) value by a factor of α *at any point in time* occurs with probability at most $1/\alpha$. The core idea of sequential testing with NSMs is to

construct a test statistic, itself a stochastic process, that is (upper bounded by) an NSM with initial (expected) value less than or equal to 1 under the null hypothesis. Ville's inequality then allows us to reject the null at confidence level $1-\alpha$ if we observe the sequence *at any point in time* to exceed $1/\alpha$.

2.4.2. E-processes

An e-process is a generalization of non-negative super-martingales similarly suitable for hypothesis testing. Concretely, for a set of distributions $\mathcal{P}$, a $\mathcal{P}$-safe e-process is a non-negative stochastic process E_t with the following property;

$$\sup_{P\in\mathcal{P}} \sup_{\tau} \mathbb{E}_P[E_\tau] \leq 1 \tag{7}$$

where the inner supremum is taken over all possible stopping times τ. A stopping time is any time that can be selected using *only* information available up to that point in time, that is, the decision to stop at time t cannot depend on data collected after time t. This property implies, via Markov's inequality, the stopping time defined as the earliest time t such that $E_t \geq 1/\alpha$ results in a level α sequential test against the null $\mathcal{P}$ (Ramdas et al., 2022). The above property guarantees that stopping an e-process at any valid stopping time results in a valid sequential test statistic.

E-processes are related to the more recognizable p-value via simple inversion. For any $\mathcal{P}$-safe e-process E_t, the sequence $p_t := \inf_{s\leq t} 1/E_s$ is an any-time valid $\mathcal{P}$-safe p-process (Grünwald et al., 2020). As its name suggests, a $\mathcal{P}$-safe p-process produces p-values against the null $\mathcal{P}$ at every stopping time.

Intuitively, e-processes can be thought of as a generalization of the likelihood ratio to handle a composite null and composite alternative. For a single null P and single alternative Q, Wald (1945) develops the sequential likelihood ratio test:

$$L_t := \frac{\ell_Q(X_1,\dots,X_t)}{\ell_P(X_1,\dots,X_t)} \tag{8}$$

The ratio measures the weight of evidence that the data was generated by Q rather than P. When the true data-generating distribution is P, the expectation of the P likelihood function evaluated at the data will be at least as great as the expectation of the Q likelihood function. In other words, $\mathbb{E}_P[L_t] \leq 1$. Here, L_t is also a NSM under P.

In general, e-processes are intimately related to test NSMs. When $\mathcal{P}$ is a singleton every NSM is also an e-process. For general sets, $\mathcal{P}$, that is when the null is composite, there may not exist a single process that is an NSM for every $P \in \mathcal{P}$. However, as Ramdas et al. (2020) show, every $\mathcal{P}$-safe e-process, E_t, is dominated by a $\mathcal{P}$-safe e-process of the form

$$E_t \leq E_t^* := \inf_{P\in\mathcal{P}} M_t^P \tag{9}$$

where M_t^P is a non-negative martingale on P with initial expectation (at most) 1. Neither E_t nor E_t^* is itself a non-negative (super-)martingale but it is upper bounded by one so Ville's inequality is still applicable.

Ramdas et al. (2022) demonstrate one particularly useful and simple construction of e-processes for a composite null and a composite non-parametric alternative which we leverage in this work. Inspired by Wasserman et al. (2020), they define the following $\mathcal{P}$-safe e-process;

$$\inf_{P\in\mathcal{P}} \frac{\prod_{s=1}^t g_s(X_s)}{\ell_P(X_1,\dots,X_t)} = \frac{\prod_{s=1}^t g_s(X_s)}{\sup_{P\in\mathcal{P}} \ell_P(X_1,\dots,X_t)} \tag{10}$$

where g_s is any "non-anticipating" probability mass (or density) function. The function g_s is "non-anticipating" if it depends only on the first $s-1$ data points. As a result, g_s can be constructed as data is being collected, learning from past data in any way it sees fit. This construction opens the door for the direct application of machine learning methods in sequential testing. In the language of Shafer and Vovk (2019), g_s is a "bet" against the null which, we as the statistician, are free to generate. Of course, the choice of g_s implicitly defines an alternative and therefore the power of the test. We can use this to our advantage to tune the sensitivity of the test toward preferred regions of the alternative space.

3. Methods

3.1. Inferring service-related stops from GPS trajectories

Suppose we have a sequence of n GPS pings that have been classified as coming from a stationary vehicle for the same driver on a shift ordered by timestamp. We are interested in grouping these pings into *service stops* representing a time and location for which a vehicle may have been involved in a service-related activity, for example a pickup or drop off.

We will do so by estimating the following random variables:

- $S_i = 1$ if the ith still ping was a service stop and 0 otherwise, and
- $C_i = 1$ if still pings i and $i+1$ were associated with the *same* service stop.

Taken together these two indicators are used to transform a sequence of still GPS pings into a sequence of service stops using the procedure given by Algorithm 1 in the appendix. We describe each still ping by three features:

1. The speed measurement of the ping denoted by v_i,

2. The time difference between pings i and $i+1$ denoted by t_i, and
3. The distance between pings i and $i+1$ denoted by d_i.

First we consider S_i. Intuitively we are more confident that a still ping is in fact a stop if the speed measurement is small. We therefore model the probability that $S_i = 1$ given a speed measurement v by the logistic function whose mean $\bar{v}$ we treat as a free parameter;

$$\Pr[S_i = 1 \mid V_i = v] = 1 - \text{LogisticCDF}(v; \bar{v}) \tag{11}$$

where,

$$\text{LogisticCDF}(v; \bar{v}) = \frac{1}{1 + \exp(-(v - \bar{v}))} \tag{12}$$

The use of the CDF of the logistic distribution captures our intuition that larger speed measurements imply that it is less likely that the ping was produced by a vehicle at a service stop. The parameter $\bar{v}$ is the median of the logistic distribution and thus should be selected as the "break-even" speed: the speed measurement at which the vehicle is equally likely to be stationary or moving. Equivalently, $\bar{v}$ is the largest (smallest) speed for which we would say that it is most likely that the vehicle was stationary (moving).

Next we consider C_i. If one or both of pings i and $i+1$ are not service stops then clearly pings i and $i+1$ cannot be associated with the same service stop. Conditioned on the event the $S_i \cdot S_{i+1} = 1$ then, the probability that $C_i = 1$ depends on the distance between the two pings, which we denote by D_i and the time difference between the two pings, denoted by T_i. The larger both D_i and T_i are the less likely it is that the two stops refer to the same location. As above, we encode this intuition via the CDF of the logistic distribution;

$$\Pr[C_i = 1 \mid T_i = t, D_i = d, S_i S_{i+1} = 1] = 1 - \text{LogisticCDF}(\rho t + d; \rho \bar{t} + \bar{d}) \tag{13}$$

where ρ serves to convert time into distance. As before $\bar{t}$ and $\bar{d}$ should be chosen to be our "break-even" time and distance: the largest time interval and distance we would be willing to accept between two consecutive pings associated with the same service stop. The conversion factor ρ has a ready interpretation as an average travel speed.

The location of a service stop is defined as the centroid of the GPS pings associated with the service stop. The timestamp of a service stop is defined as the earliest timestamp of the GPS pings associated with the service stop.

3.2. Non-stationary binary sequence testing

We observe a sequence of binary-valued random variables $\{X_i\}_{i=1,\dots,T}$. In this setting we imagine that they represent a reward or penalty experienced by an agent interacting with its environment, but this construction serves merely to motivate our particular application. Our question is to determine the extent to which they appear to be learning from experience: in short we ask, is the probability of $X_i = 1$ increasing (decreasing) over time?

In order to detect an increasing or decreasing probability of reward over time we will construct a generalization of a likelihood ratio known as an e-process against the null that the sequence of binary variables are i.i.d. Bernoulli random variables with a common unknown mean. Following Ramdas et al. (2022), our process will take the form;

$$E_t := \frac{\prod_{s=1}^{t} (\hat{p}_s)^{X_s} (1 - \hat{p}_s)^{1 - X_s}}{\left(\frac{n_1^t}{t}\right)^{n_1^t} \left(\frac{n_0^t}{t}\right)^{n_0^t}} \tag{14}$$

where n_1^t (n_0^t) are the number of times a 1 (0) has been observed up to and including time t and $\hat{p}_s \in [0, 1]$ is any estimate of $\Pr[X_s = 1 \mid X_1, \dots, X_{s-1}]$. This estimate may depend on the previously observed data $X_1, \dots, X_{s-1}$. The process by which the estimates $\hat{p}_s$ are constructed will affect which alternatives the test has power under. In this work we will focus on three related estimators:

$$p_t^{\pm} = \frac{\sum_{s=1}^{t-1} (1 - \gamma)^{s-1} X_{t-(s-1)}}{\sum_{s=1}^{t-1} (1 - \gamma)^{s-1}} \tag{15}$$

$$p_t^{+} = \max\left\{ \frac{n_1^{t-1}}{t-1}, p_t^{\pm} \right\} \tag{16}$$

$$p_t^{-} = \min\left\{ \frac{n_1^{t-1}}{t-1}, p_t^{\pm} \right\} \tag{17}$$

where $\gamma \in (0, 1)$ can be specified in advance or learned as data is collected. The first estimator, $p_t^{\pm}$, is simply an exponential weighted average of the data collected so far, representing the non-parametric alternative that the data-generating process is non-stationary but uses similar probabilities for data drawn nearby in time. A simple way to construct an estimator that makes Eq. (14) sensitive in particular to non-stationary Bernoulli sequences for which the event probability is increasing (decreasing) is to leverage the following observations. When the probability of $X_t = 1$ is increasing (decreasing) in time, then the exponential weighted mean, $p_t^{\pm}$ will be:

1. a better estimate of $\mathbb{E}_{t-1}[X_t]$ than the empirical mean, and
2. larger (smaller) than the empirical mean.

The second and third estimators encode these observations by restricting the range of the exponential weighted mean. The second estimator, p_t^+, takes on values that are at least as large as the empirical mean and represents the alternative that the data-generating process has increasing probability of $X_t = 1$ over time. The third estimator, p_t^-, takes on values that are no larger than the empirical mean and represents the alternative that the data-generating process is decreasing the probability of $X_t = 1$ over time. We note that the validity, that is the type-I error, of the test is unaffected by our choice of alternative likelihood.

Under the null, the likelihood using $p_t^\pm$ will, in expectation, be at most the likelihood using the empirical mean precisely because the empirical mean is the maximum likelihood estimate for the mean of a sequence of i.i.d. Bernoulli variables. As a result, we have $\mathbb{E}_{t-1}[E_t] \leq 1$. However, under their respective alternatives and using the appropriate value of γ, $p_t^\pm$, p_t^+, and, p_t^- provide a better estimate of $\mathbb{E}_{t-1}[X_t = 1]$ than the empirical mean does; as a result we should expect $\mathbb{E}_{t-1}[E_t] > 1$.

Although p_t^+ has power under increasing binary sequences, it does not have power under decreasing sequences. When we test a *decreasing* sequence using p_t^+, the exponential weighted average will fall below the empirical mean and the max will cause $p_t^+ = n_1^{t-1}/t-1$, and we have $\mathbb{E}_{t-1}[E_t]$ smaller than, but close to, 1. Conversely, p_t^- does not have power under increasing sequences. As a result, we may construct a "two-sided" test via two applications of the test, the first against the increasing alternative and the second against the decreasing alternative.

In short, we are interested in testing the null H_0 against two non-parametric alternatives H^+ and H^-:

H_0 The sequence is i.i.d. Bernoulli with unknown mean
H^+ The sequence is drawn from a Bernoulli with increasing mean
H^- The sequence is drawn from a Bernoulli with decreasing mean

In our context the observations are events that have occurred as an artefact of some agent's decision process. When we reject the null in favor of H^+ we say that the agent is *event-seeking*. When we reject the null in favor of H^- we say that the agent is *event-avoiding*.

Our question may be considered as a *changepoint detection problem* and our test as an *online changepoint detection* procedure. Given a sequence of observations $\{X_t\}_{t=1,2,\ldots}$, changepoint detection aims to identify the time point, ν, at which the data-generating distribution changes from one given family of distributions, $\mathcal{P}$, to another given family of distributions $\mathcal{Q}$ while controlling the probability of false detection (Shin et al., 2022). Concretely we have,

$$X_0, \ldots, X_\nu \sim P \in \mathcal{P} \tag{18}$$

$$X_{\nu+1}, \ldots \sim Q \in \mathcal{Q} \tag{19}$$

It is important to note that in the most general setting $\mathcal{P}$ and $\mathcal{Q}$ are sets of distributions over *sequences*.

In the language of changepoint detection our question becomes: at what time does the sequence of observations start increasing? Our set of pre-change distributions $\mathcal{P}$ is the set i.i.d. Bernoulli sequences and our set of post-change distributions $\mathcal{Q}$ is the set of distributions over sequences whose expectations are increasing (decreasing) in time. Our statistic E_t will identify the point in time at which enough evidence has been gathered to say that the sequence is not i.i.d Bernoulli and may be increasing.

4. Data

4.1. Ride-sourcing driver GPS trajectories

We source our GPS data from the "rideshare and delivery assistant" app Gridwise.[1] The Gridwise GPS data contains **183,548,522** individual GPS pings from **5352 drivers** in the Pittsburgh region for two **nine-month time periods**: **March to November of 2019** and **March to November of 2020**.

The GPS data is also labeled with an "activity type". The activity type is a device-generated variable which reports the physical activity of the device inferred from its kinematic sensors. It is a categorical variable taking one of seven values summarized in Table 1. Gridwise drivers spend most of their recorded time driving, a finding consistent with the use case for the Gridwise app. Of particular interest are the GPS pings classified as 'still'. This set of pings contains the stops made by the driver and will form the basis of our subsequent analysis.

The GPS trajectories themselves are of high quality. Over 90% of pings are within 30 s of the most recent ping from the same driver and three-quarters are within 6 s. Despite the fact that drivers are logging on and off the platform over weeks if not months, the high temporal density of pings implies that the data contains long uninterrupted sequences of GPS pings. The geographic coordinates of half of all pings lie in a reported confidence radius of 5 meters (roughly 1 car length); 90% lie in a reported confidence radius of 16 m.

The GPS trajectories are highly concentrated spatially. Of the US census tracts containing at least one GPS ping, the top 1% account for approximately 18% of all GPS pings. The tracts containing the Pittsburgh International Airport and Pittsburgh's Downtown areas alone account for about 11% of all GPS pings.

[1] https://gridwise.io/

Table 1
Number and percent of GPS pings by activity type.

Activity type	n	%
unknown	5,697,293	3.10%
still	11,461,336	6.24%
on_foot	5,842,891	3.18%
walking	1,422,662	0.78%
running	93,625	0.05%
on_bicycle	286,545	0.16%
in_vehicle	158,744,169	86.49%

4.2. OpenStreetMaps feature tags

OpenStreetMaps provides open source geographic data for the entire world, maintained by a community of users (OpenStreetMap contributors, 2022). OpenStreetMaps represents physical entities in the world by associating one or more tags to a geographic shape. A tag is a ':' delimited key–value pair of text data (e.g. `amenity:restaurant`). OpenSteetMaps uses a "free tagging system" meaning that map contributors are able to use any text they wish when tagging a geographic shape with data. However, informal standards have been adopted within the community around the most commonly used tags.[2]

Our aim in using OpenStreetMap feature tags is to give context to the driver GPS trajectories and, ultimately, to find tags that were *relevant* to the driver's decision-making process. To that end we focus on tags with the following keys:

- aeroway: airport-related tags.
- amenity: "useful and important facilities" for people.
- building: physical buildings and their associated use (e.g. residential, commercial, office).

In total, there are 31 tags of interest within the above tag keys.

We also leverage the `opening_hours` tag which reports the days and times when the entity is open. In particular this value is used to establish whether or not an entity was open at a particular moment (e.g. when a driver made a stop nearby).

5. TNC driver GPS analysis

There are several notable differences in the GPS data between 2019 and 2020 which we attribute to the onset of the COVID-19 pandemic. Our analysis suggests that in the months following the onset of the COVID-19 pandemic,

1. Compared to the same period the year prior, there were fewer ride-hailing drivers who took on an increased number of shorter shifts, and,
2. driver activity increased in the evening and decreased during the day.

An important caveat for the following analysis is that the sample of drivers represented on the Gridwise platform cannot be considered a random sample of TNC drivers. As a result, we cannot immediately use this analysis to conclude anything about the population of ride-hailing and delivery drivers. Drivers self-select to use Gridwise, presumably because they believe it will enable them to increase their earnings. Nevertheless, Gridwise data provides an unparalleled detailed look into the behavior of ride-hailing and delivery drivers.

From March to November 2019, Gridwise saw a roughly 50% increase in monthly active drivers on the platform. After the onset of the COVID-19 pandemic, the platform steadily lost drivers. By November of 2020, the platform saw over 60% fewer drivers than it did in March of 2020. The decrease in monthly active drivers on the Gridwise platform is consistent with decreases in platform utilization of ride-hailing services more broadly during this time period. In their second quarter filing with the U.S. Securities and Exchange Commission (SEC), Uber reported a 66% decrease in net revenue from mobility services and a 162% increase in net revenue from delivery services relative to the second quarter of 2019. Across all segments, Uber saw a 44% decline in monthly active customers and a 55% decrease in trips compared to the second quarter of 2019 (Uber Technologies, 2020).

However, analysis of the monthly active drivers on Gridwise paints a different picture: instead of a sudden decline followed by slow recovery of drivers, Gridwise saw a steady decline in monthly active drivers over time. Drivers on Gridwise in 2020 had longer platform lifetimes than drivers in 2019 did, and spent more time on the platform on a per-driver basis. Measured as the number of days elapsed between the earliest and latest GPS pings associated with a driver id, average driver platform lifetime increased by over 30% from 2019 to 2020. Measured as the number of pings produced per driver, driver time on platform increased by nearly 30% from 2019 to 2020 and nearly 65% from July to November of 2020 compared to the same period in 2019. Taken together, we speculate that as the earning potential of being an TNC driver declined after the onset of COVID-19, the drivers that remained on the Gridwise platform were those who relied on TNCs for as large portion of their income. While drivers who had other options or only casually drove for TNCs simply left the platform, those that stayed doubled down, driving more shifts and longer and later hours.

[2] See https://wiki.openstreetmap.org/wiki/Map_features.

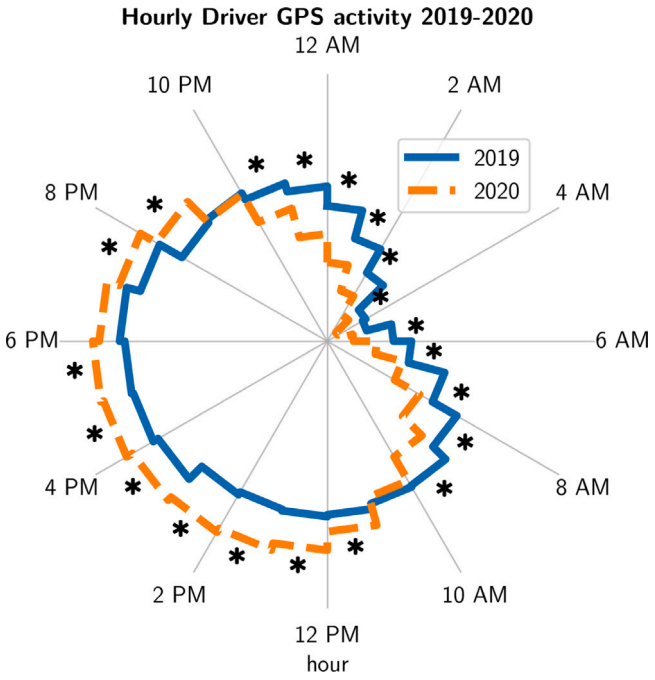

Fig. 1. Histogram of GPS pings by hour of day. Asterisks denote hours with statistically significant (Bonferroni corrected p-values at a 99% confidence level) difference between the histograms.

Despite the decrease in drivers, the number of shifts on Gridwise remained largely the same, but shifts tended to be shorter with later start and end times. Drivers in 2020 tended to start their shifts later than their 2019 counterparts: comparatively more shifts started between 2pm and 10pm and fewer between 1am and 1pm. These shifts ended later too: the hours between 5pm and 3am saw an increase in share of shifts ending from 2019 to 2020. In total, during the initial stages of COVID-19, drivers were more active in the afternoon and evening (11am–9 pm) and less active in the late-night and mornings (9pm–11am) compared to 2020. Each day, defined as the 24 h period starting at midnight local time, we computed a histogram of pings using 24 hour-aligned bins, each bin reporting the fraction of pings that occurred in each hour. We then applied a Bonferonni-corrected Welch's t-test to determine the hours for which the average fraction of daily pings changed from 2019 to 2020. We report the average fraction of pings for 2019 and 2020 and significance at the 99% confidence level in Fig. 1

5.1. OpenStreetMap analysis

What were drivers doing differently in 2020 compared to 2019? To address this question we join still GPS pings from the Gridwise data with OpenStreetMaps entity tag data. We define a ping "interaction" with a tag as the event that an entity with the given tag was within 250 ft of the given ping. The interaction is a many-to-many relationship between entities and pings: each ping can have interactions with multiple entities (some having the same tag) and each entity may have interacted with multiple pings. We then sum the interactions for each tag over each hour and normalize by the total number of interactions in that hour, giving us a measure of collective driver interest in each tag by hour of day. In other words, the fraction of interactions answers the question: of all the entities nearby driver stops, what fraction of them had a given tag?

One of the more immediate impacts of COVID-19 on TNC ride-hailing operations was the reduction in airport trips due to the reduction in air travel. Airport trips represented 15% of Uber's 2019 rides gross bookings (Uber Technologies, 2019). In April 2020 Pittsburgh International Airport saw a 96% decrease in total passengers compared to April 2019 (Allegheny County Airport Authority, 2020). We find that TNC drivers had dramatically fewer interactions with the airport and with hotels in 2020 as compared to the same month in 2019, both in absolute interactions and as a fraction of total interactions, as shown in Fig. 2. Remarkably, the fraction of airport interactions tracks very closely with the passenger volumes (black dashed line in Fig. 2) at the Pittsburgh airport from March to November of 2020.

A second noticeable effect of COVID-19 was that demand for TNC services shifted from passenger rides to delivery. In Uber's 2020 s quarter (Q2) filing, passenger rides decreased substantially while trips on Uber's food delivery platform soared (Uber Technologies, 2020). However, total activity, measured as both the number of trips (rides and delivery) and gross bookings, decreased. Many restaurants experienced a similar trend: an increase in delivery offset by a decrease in dine-in patrons. In their Q2 2020 filing, McDonald's reported a 31% decrease in sales and a decrease in guest count compared to Q2 2019 even as delivery

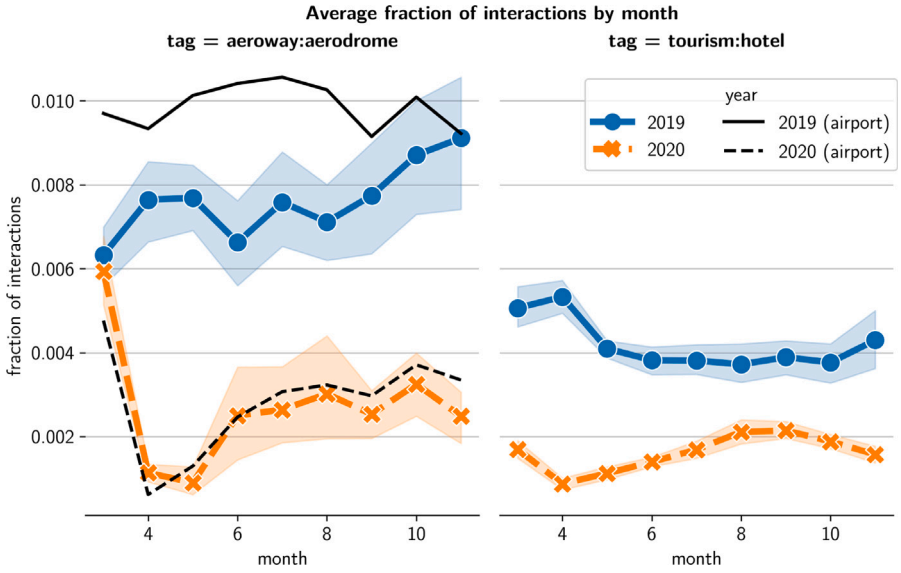

Fig. 2. Average fraction of TNC driver interactions with tourism-related tags by month in 2019 and 2020. For comparison, Pittsburgh international airport passenger volumes (2019 volumes in solid black, 2020 volumes in dashed black) have been re-scaled and super-imposed onto the airport interactions.

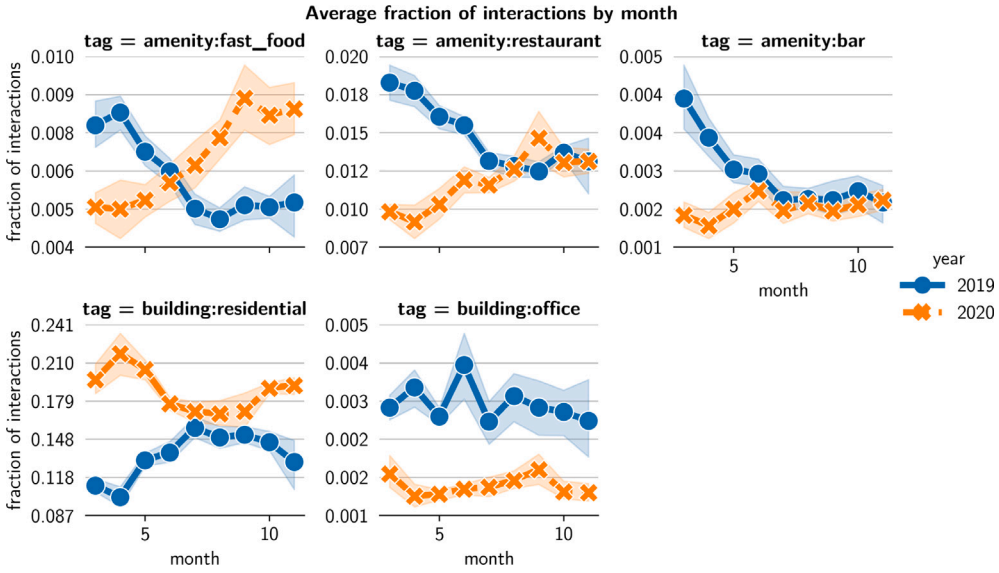

Fig. 3. Average fraction of TNC driver interactions with passenger and delivery-related tags by month in 2019 and 2020.

sales increased "significantly" (McDonald's Corporation, 2020). It stands to reason that any gains in TNC activity restaurants saw from Uber Eats were, on average, offset by even greater losses in TNC activity from passenger rides.

Can we demonstrate changes in tags that would be consistent with rides vs eats changes? Specifically we will focus on driver interactions with entities having the `amenity:fast_food`, `amenity:restaurant`, and `building:residential` tags. To ground our analysis we will also look at the `amenity:bar` and `building:office` tags. Under the ride-hailing prominent pre-COVID conditions we might expect bars and restaurants to have similar interactions with TNC drivers; likewise we might expect residences and offices to have similar interactions. Under post-covid conditions, when food delivery increased dramatically and passenger trips decreased dramatically, we should see diverging interaction trends between bars and restaurants and between residences and offices. Lastly, as Fig. 1 demonstrates we expect to see relative more activity in 2020 during the period from 11am until 9pm and relatively less activity during the period from 9pm until 11am. We term the periods the "late" and "early" periods respectively and will examine interactions in these two periods separately. Our findings are shown in Fig. 3.

Overall, we find that tag incidence data is consistent with the notion that COVID-19 produced a shift in consumer behavior away from travel and towards delivery. We find that the fraction of fast food and restaurant interactions increased from March

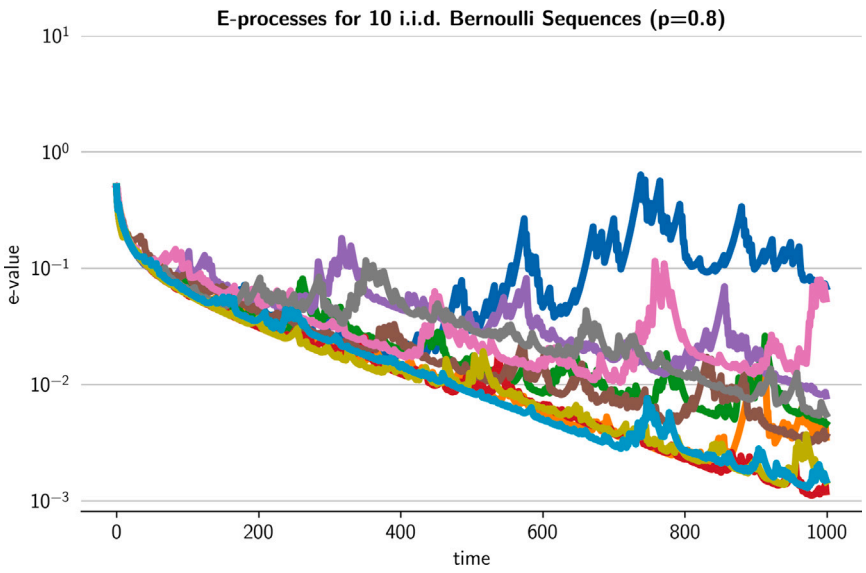

Fig. 4. Our e-processes demonstrated on 10 different i.i.d Bernoulli sequences ($p = 0.8$).

to November 2020 despite a decreasing trend in those same months in 2019. During the late period, the fraction of fast food and restaurants interactions in 2020 met or exceeded the same-month fraction in 2019. In contrast, the fraction of bar interactions in 2020 was flat and below the 2019 same-month fraction. The fraction of residential interactions in 2020 comfortably exceed the same-month fractions in 2019 for both the early and late periods. In contrast, the fraction of office interactions in 2020 remain decidedly below the same-month fractions in 2019 for both periods.

While it is tempting to read into these figures, it remains challenging to attribute these observations to a particular cause. As noted previously restaurant interactions are due a mix of delivery and passenger rides, both of which we know to have changed over time. Similarly, residential interactions are a mix of home delivery and home rides. Since many more people were home for greater amounts of time, it is not clear how much of the difference in residential interactions is attributable to increased food delivery as opposed to more time spent at, and therefore traveling to or from, home.

6. Experiments

6.1. Simulated data

We demonstrate our e-process on simulated data. First, we consider sequences of i.i.d. Bernoulli data on which our e-process should fail to reject the null. Second, we consider sequences of rewards generated by agents interacting with a k-armed Bernoulli bandit where we expect our e-process should reject the null.

6.1.1. Stationary Bernoulli

We first apply our e-process to i.i.d Bernoulli data ($p = 0.8$). The values of the 10 e-processes are shown in Fig. 4. As expected, none accumulates enough evidence for us to reject the null; here we use the increasing alternative to ease comparison with the following section. Though we fail to reject the null, the e-process does not stray too far from one, yielding e-values between 10^{-3} and 10^{-1} after 1000 observations. By comparison, the likelihood ratio of the point alternative $p = 0.7$ against the (ground truth) point null of $p = 0.8$ yields a e-value around 10^{-8} after 1000 observations. In short, the exponential weighted mean is a poorer estimate of the true probability than the empirical mean, but not by a wide margin. This is helpful particularly in the changepoint application because if the data-generating distribution changed after $t = 1000$ it should not take long for the e-process to rebound and detect the change.

6.1.2. Bernoulli bandit

We create a k-armed bandit ($k = 16$) with binary ($\{0, 1\}$) rewards whose arm means (p_i for $i = 1, \ldots, 16$) are drawn i.i.d from a uniform distribution. Each time an agent pulls an arm it receives a reward of 1 with probability p_i. We examine the behavior of 10 agents each of which employs Thompson Sampling with a Beta(1, 1) prior (Agrawal and Goyal, 2012). Each agent interacts with the same bandit instance for 1000 iterations. The behavior of the agents and their respective e-processes may be contrasted with the e-process for the i.i.d. Bernoulli sequences above. When an agent employs a fixed policy, the resulting reward sequence is i.i.d. Bernoulli and our test will fail to reject the null.

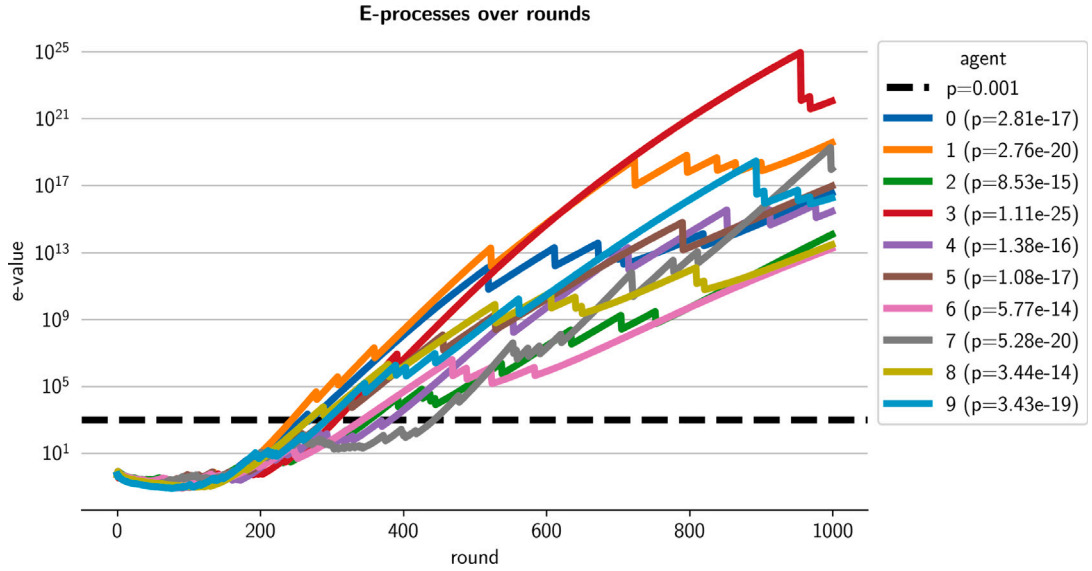

Fig. 5. Our e-process over bandit rounds, log scale. Black line indicates the 99.9% confidence level.

One e-process with an increasing alternative is generated for each of the reward sequences and shown in Fig. 5. The legend lists the p-value derived from the test statistic next to its associated agent. For each of the agents, the e-process accumulates enough evidence to soundly reject the null. When we instead generate the e-process with the decreasing alternative, we fail to reject the null: the exponential weighted mean falls below the empirical mean and the alternative likelihood must use the empirical mean.

In short, this use case demonstrates the effectiveness of the test in establishing whether an agent *could possibly* have learned something about the environment they are interacting with, without knowledge of the action space or the environment.

6.2. Ride-hail driver GPS trajectories

We now apply our methodology to the GPS trajectories of ride-sourcing drivers. Specifically we analyze the on-shift GPS trajectories of drivers from March through November 2020. We apply the probabilistic model described in Section 3.1 to pings classified with the "still" activity type. We set the "break-even" speed, time, and distance ($\bar{v}, \bar{t}, \bar{d}$, respectively) to 8 miles per hour, 5 min, and 200 ft. The conversion factor from time to distance, ρ, was set to the average measured speed over all trajectories in the 2019 dataset: 14.48 meters per second (roughly 32 miles per hour). These values were assumed and heuristically validated against random subsets of the 2019 data by visual inspection of the stop inference procedure labels on a map. For the purposes of the test, we ignore all drivers who have fewer than 500 service stops; leaving 106 drivers. Each service stop is associated with a binary vector of length equal to the number of tags of interest as identified in Section 4.2. Non-zero entries in this vector represent the presence of at least one OpenStreetMaps entity with the given feature tag that is open at the time of the service stop and within 45 meters (roughly 150 ft) of the location of the service stop. For entities with no opening hours information, the entity was considered open if at least half of the entities having the same tag with opening hours information were open at that time. If no entity having the same tag was labeled with opening hours information, the entity was assumed to be open at all times.

We perform one two-sided test against the sequence of tag indicators over service stops for each tag and each driver, yielding 3,286 (106 drivers by 31 tags) individual tests. Within the exponential weighted mean function in the alternative likelihood we set the weight parameter γ as $1 - \exp(-\ln(2)/64)$, which ensures that the weight of each sample reduces by half every 64 service stops. This value was chosen as it seemed to provide a good balance of bias and variance in tests on the simulated data generated in Section 6.1. As we previously noted, more complicated schemes could attempt to learn this value from data during the test.

The two-sided test combines two runs of the learning test, the first uses the increasing alternative and the second uses the decreasing alternative. In effect, the first test supposes the agent wishes to maximize the tag incidence over time and the second supposes that the agent wishes to *minimize* the tag incidence over time. In our setting, when we reject the null using the increasing alternative we identify potential *tag-seeking* behavior and when we reject the null using the decreasing alternative we identify potential *tag-avoiding* behavior.

We also test the behavior of the drivers as a collective. To do so, we simply combine the service stops from all drivers into a single sequence of service stops in chronological order. To construct this sequence we use service stops from all 571 drivers resulting in 197,920 service stops. To account for the greater temporal density of service stops we set the weight parameter γ to produce a half-life of roughly one week. One two-sided test is applied to each tag from this single stream of service stops.

Bonferroni correction (Shaffer, 1995) is applied to the p values returned by the test. A driver tag pair is deemed significant if its p value, multiplied by the number of driver-tag pairs tested, is less than $(1-\alpha)/2$ for any confidence level $\alpha \in (0, 1)$. In this paper we take $\alpha = 0.95$.

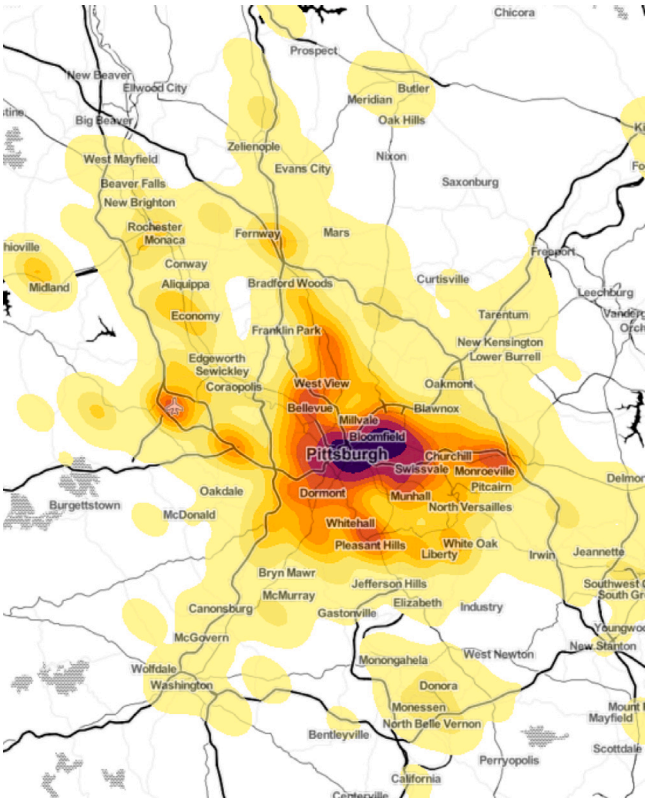

Fig. 6. Kernel density map of service stops.

7. Results

7.1. Inferred service stops

Our inference process distills 2,552,460 still GPS pings from 648 drivers into 233,349 service stops from 571 drivers. On average service stops last about 12 min and contain 3 GPS pings. Service stops in our data range over 3235 square miles of Southwestern Pennsylvania intersecting six counties. The bulk of the activity, however, occurs in the Pittsburgh area and its containing county. A heatmap of all service stops is shown in Fig. 6 whose extent has been narrowed for visual clarity. Service stops are heavily concentrated in the urban core of Pittsburgh, with substantial density across its inner suburbs.

7.2. Statistical hypothesis test results

At the 0.95 confidence level, after Bonferroni correction, we find 62 significant driver-tag pairs exhibiting tag-seeking behavior and 55 significant driver-tag pairs exhibiting tag-avoiding behavior out of the 3,286 driver-tag pairs tested.

Tables 2 and 3 report the Bonferroni-corrected p-values for tag-seeking and tag-avoiding behavior respectively. In both tables, the driver identifiers have been elided and tags which appeared in more than one significant driver-tag pair are listed once alongside the minimum p-value for that tag over drivers and the number of significant driver-tag pairs the tag appears in (column n). An extended summary of test results is given in Table B.4 in Appendix B.

We now turn our attention to the tests performed on the sequence of service stops aggregated over all drivers. In this setting tag-seeking and tag-avoiding behavior should be taken to refer to the drivers as a collective. At the 0.95 confidence level, after Bonferroni correction, we find 13 significant tags exhibiting tag-seeking behavior and 8 significant tags exhibiting tag-avoiding behavior. Tables B.5 and B.6 in Appendix B report the Bonferroni-corrected p-values for tag-seeking and tag-avoiding behavior respectively across the drivers as a collective as well as the timestamp at which the detection was made.

Notably, our procedure identifies the moment when enough evidence was gathered to conclude that the tag incidence probability was not constant in time and likely either increasing or decreasing. Conforming to our earlier analysis of airport trips we find that driver were likely decreasing their airport trips as early as late March and began slowly increasing again some time before early September. In contrast, food delivery related tags began to increase, starting with fast food and commercial areas in mid April followed by residences and restaurants in early May. It was comparatively longer before increases were seen in passenger oriented

Table 2
Bonferroni-corrected p-values for tags from driver-tag pairs exhibiting tag-seeking behavior.

Tag	p value	n
amenity:restaurant	9.43e−158	4
building:commercial	8.43e−127	8
amenity:college	5.32e−61	1
aeroway:aerodrome	1.30e−58	10
amenity:shelter	4.15e−38	5
building:residential	8.12e−38	17
amenity:place_of_worship	1.71e−15	1
amenity:grave_yard	5.06e−12	1
amenity:fast_food	8.07e−08	4
amenity:casino	1.75e−06	1
amenity:library	2.83e−06	1
amenity:university	4.29e−05	1
amenity:doctors	4.60e−03	1
amenity:dentist	7.41e−03	1

Table 3
Bonferroni-corrected p-values for tags from driver-tag pairs exhibiting tag-avoiding behavior.

Tag	p value	n
aeroway:aerodrome	2.94e−62	17
building:residential	1.11e−42	15
building:commercial	1.69e−32	5
amenity:restaurant	4.11e−17	4
amenity:shelter	6.46e−15	3
amenity:fast_food	1.44e−14	7
amenity:university	1.71e−10	1
amenity:casino	2.33e−08	1
amenity:cafe	5.40e−05	1
building:office	7.08e−03	1
amenity:hospital	8.10e−03	1

tags: theaters showed increases by June, casinos in August, bars in late October, and offices in November. Colleges and Universities showed increases by early September, which could be attributed to the beginning of the Fall semester, and highlights the difficulty in using this method to explain an increase in frequency rather than simply identify it.

8. Discussion

When we reject the null for a driver tag pair, we are saying that the probability of observing a service stop near an open entity with that tag is unlikely to be constant in time. Further, by applying the two-sided test, we can identify driver-tag pairs in which the tag probability could possibly be increasing over time, and those for which the tag probability could possibly be decreasing over time.

Figs. 7 and 8 demonstrate the running average tag incidence for the top 10 most significant driver-tag pairs over time. The 'reward signal' in our case is a binary-valued variable representing whether or not the service stop was nearby an open entity with a given tag. As a result, the time-averaged reward is simply the frequency with which the tag occurs nearby a service stop and is an empirical estimate of its probability. A driver-tag pair for which we reject the null in the tag-seeking test should exhibit increasing frequency in stops nearby the tag over time. Similarly, a driver-tag pair for which we reject the null in the tag-seeking test should exhibit decreasing frequency in stops nearby the tag over time. Accordingly, this behavior is observed for both the tag-seeking and tag-avoiding driver-tag pairs.

Our results suggest that our procedure is a highly effective behavioral trend detector by offering early trend detection with probabilistic guarantees. It can be observed in Fig. 7 that the tag incidences (particularly `amenity:restaurant`, `building:commercial`, and `building:residential`) increase dramatically *after* we reject the null in favor of the tag-seeking alternative. In fact, most of the increase in tag incidence occurs after the null has been rejected. The `building:commercial` tag (orange in Fig. 7) reaches a 60-day moving average low in mid-May at around 0.09 before rising over the next ten weeks to a high of 0.65 in early August. We first identify the increase in early June when we reject the null, meaning that the trend was identified within the first three weeks and first 10% of its eventual rise. Or course our test does not make any guarantees (probabilistic or otherwise) about what happens after we reject the null. Other tag incidences in Fig. 7, for example `amenity:fast_food` and `amenity:shelter`, see only a modest increase in their frequency after we reject the null. However, the sensitivity of our procedure is tuned intuitively via the confidence level which may be directly interpreted as the probability of a false alarm. Performing a single test at the α confidence level implies that a false rejection of the null (we identify a stationary sequence of tag-incidences as being tag-seeking) occurs with probability at most $1-\alpha$. The Bonferroni correction applied to our experiments further guarantees that the

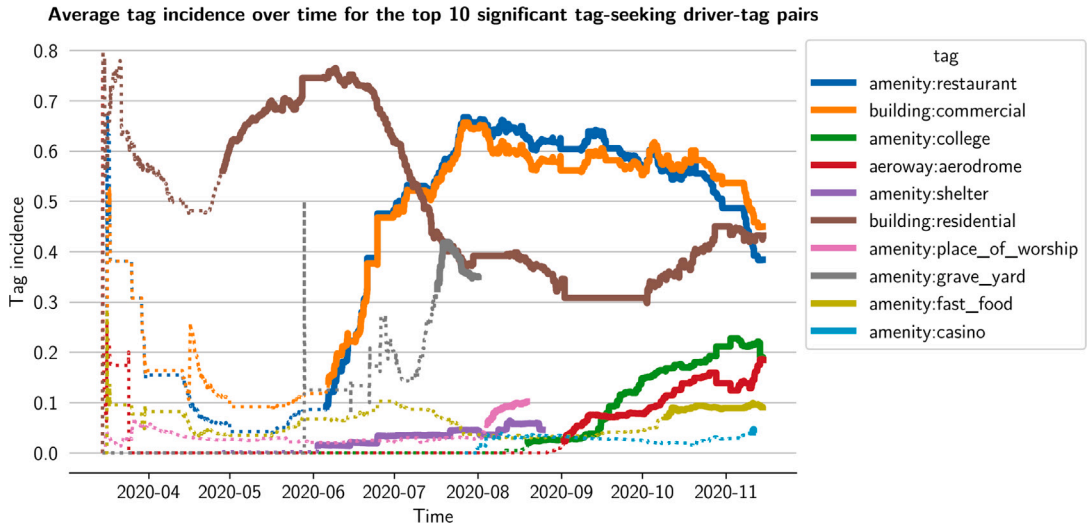

Fig. 7. 60-day moving average tag incidence over time for the top 10 significant tag-seeking driver-tag pairs. Solid portions of the line begin when the null was rejected.

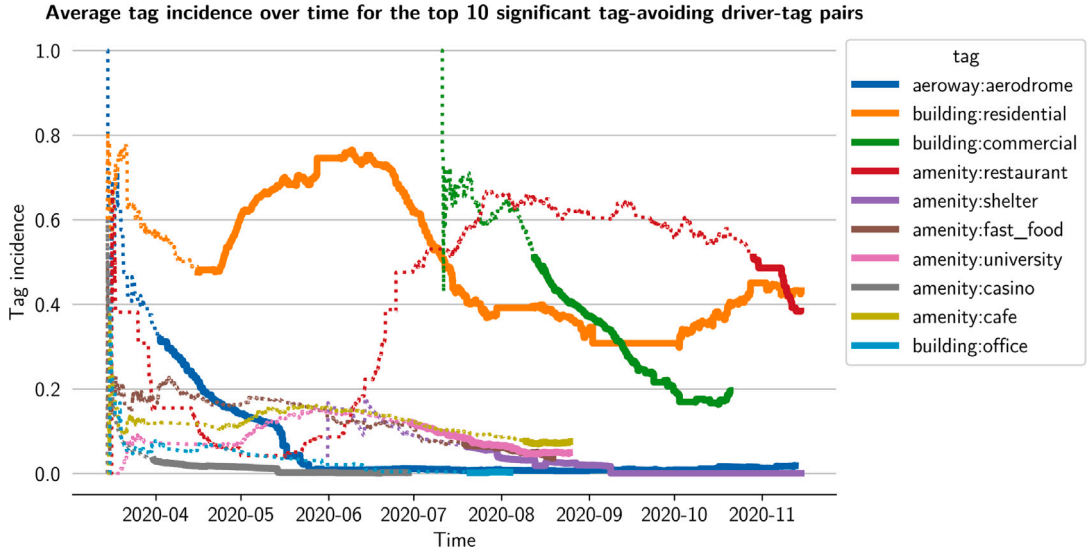

Fig. 8. 60-day moving average tag incidence over time for the top 10 significant tag-avoiding driver-tag pairs. Solid portions of the line begin when the null was rejected.

probability of *any* false rejection occurs (over all tests) is at most 5%. For use cases where false alarms are costly, higher confidence levels may be used to require a greater level of evidence to reject the null. In these settings we should expect it to take longer and more of the trend to have been observed, for tag-seeking or tag-avoiding behavior to be identified. Conversely when early detection is more important and false alarms are tolerable, the confidence level may be lowered.

One caveat of the sequential nature of this hypothesis test is that it reports if *at any point* the sequence of tag frequency *up to that point in time* appears to have increased (decreased). A sequence of average rewards which initially rises and then falls may have risen enough initially to reject the null in the tag-seeking test, and if it falls enough, reject the null in the tag-avoiding test. In fact, this is evident in our test results: the `amenity:restaurant` is significantly increasing (blue in Fig. 7) and also significantly decreasing (red in Fig. 8). This same phenomenon can be observed with the `building:commercial` and `building:residential` tags.

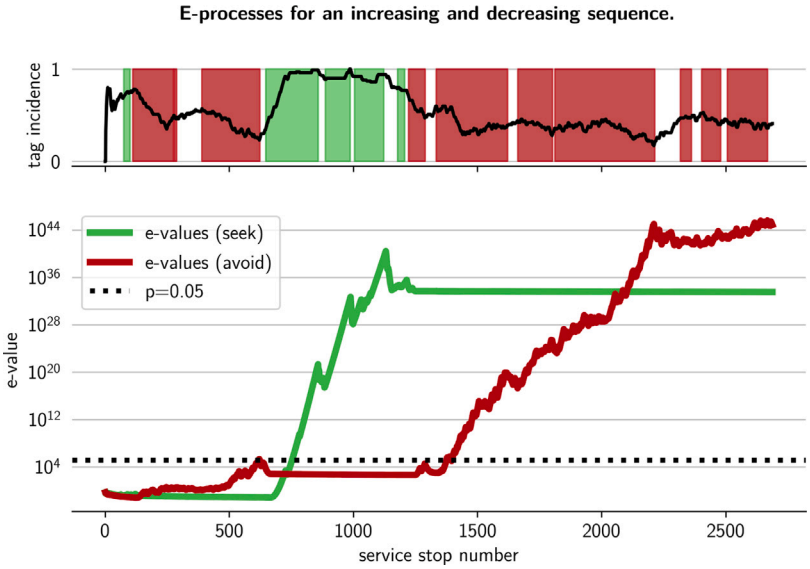

Fig. 9. How the e-processes for increasing and decreasing sequences respectively accumulate evidence.

When we inspect the e-processes for the `building:residential` driver tag pair in Figs. 7 and 8 we gain some intuition about how the e-process accumulates evidence. In Fig. 9 both the increasing alternative (seek) and decreasing alternative (avoid) e-processes are plotted in the lower panel. In the upper panel we show a moving average of the tag incidence which serves as a local estimate of the probability of stopping at a residential building. When the likelihood ratio favors the alternative over the null, the e-process will rise, indicating that it is accumulating evidence. The rate at which it rises is indicative of how much it favors the alternative. Conversely, when the likelihood ratio favors the null over the alternative the e-process will fall. Our e-process will increase much faster under the alternative than it will decrease under the null and this is evident in the figure. Under the null, the alternative likelihood will fall back to the likelihood under the empirical mean in hindsight, which will result in likelihood ratios near one.

Statistically speaking, the rise and fall of e-processes is irrelevant; once the e-process crosses our significance threshold we are justified in rejecting the null, no matter what it does after. However, qualitatively it is instructive to relate the accumulation of evidence in the e-process with the increasing and decreasing probability in the sequence. In Fig. 9 we illustrate this by qualitatively (via smoothing and finite differences) defining regions where the e-process is accumulating evidence. In the upper panel, timespans during which the e-process is accumulating evidence that the driver is avoiding residential buildings are shaded in red, while timespans during with the e-process is accumulating evidence that the driver is seeking out residential buildings is shaded in green. We emphasize that we are no longer using any statistical properties of the e-process in defining these regions so there is no statistical claim to be made regarding these regions. What this does show is that there is an intuitive alignment of the e-process and the observations: when the frequency with which we observe an event increases, the increasing alternative e-process rises and the decreasing alternative e-process stagnates or declines.

A second caveat is that the test itself cannot explain *why* the time-average rewards increase. When we are comfortable with the assumption that the environment (i.e. the action-conditioned reward distribution) is stationary, then, by process of elimination, we must conclude that it was the agent, rather than the environment that changed behavior over time. In our setting, this assumption may not hold, as COVID-related restrictions shifted considerably over the lockdown period. So what we can say in interpreting the test results is that we can identify certain drivers who over the course of the lockdown changed their behavior to either seek out services near tags that they previously had not frequented or avoid services near other tags that they previously had. It might be the case that these drivers sought or avoided these entities because of a change in government restrictions, a change in some other aspect of the environment, or because of a change in how they valued these entities.

Perhaps the most conservative interpretation of these results is that statistical significance represents a shift in the underlying process generating the driver's behavior. There are a myriad of factors involved in producing an observation that a driver made a service stop at, say, a residential building: not only the drivers own decisions but the demand for TNC services, decisions of stores on the delivery platform, the platform decisions, government regulations, seasonal trends, and even traffic incidents or road work among others. Our results can be interpreted as gathering evidence that the process is now different, without specifying which component we believe to have been driving the change.

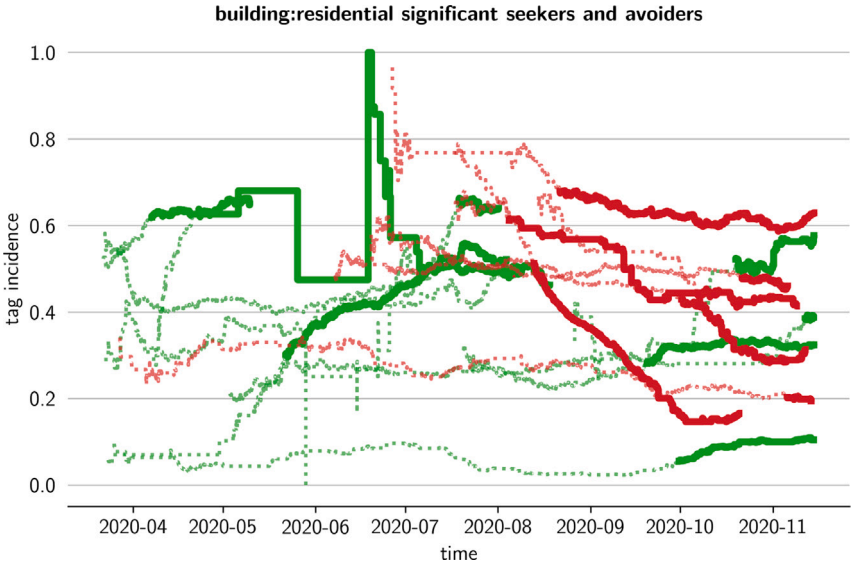

Fig. 10. Moving average tag incidence for significant seekers (green) and avoiders (red). Lines are dotted until they are found to be significant, then solid thereafter.

If we had a more specific hypothesis and additional data however, this methodology gives us the tools to help evaluate them. First, if we had more precise time series information about the specific entities in question, and in particular what COVID restrictions were applied to them at that time, we could use a conditional incidence as our sequence of observations, that is, for example, the event that a driver stopped nearby a restaurant *and* it was open for takeout. Another example would be the event that a driver stopped nearby a venue *and* there was an event ongoing at that time. At a technical level this is not different from how we used the opening hours information to condition our entity incidences.

Second, if we were to hypothesize that another binary metric was driving tag incidence, then if we applied our test to both, we should expect first that they are both significant with respect to the same alternative, and that the time at which the metric because significant precedes the time that the tag incident became significant. This application does not test the hypothesis directly and is most useful when there may be an unknown temporal lag in the relationship between the two. For example, we could hypothesize that the event that COVID hospitalization or case count were above a certain threshold would make take-out ordering more likely. In combination with the previous method, we would test the tag incidence of service stops nearby restaurants open for take out and also test the sequence of observations of COVID hospitalizations or case counts, hoping to see evidence for the latter accumulate to significance before the former.

Finally, it may be possible to directly test these hypotheses by modifying the e-process appropriately. In this setting we wish to test against the null that the two metrics are independent where the alternative is that they are dependent, potentially constraining the estimates to specifying the direction of the relationship (e.g. for binary Y, $\mathbb{E}_{t-1}[X_t \mid Y_t = 1] \geq \mathbb{E}_{t-1}[X_t \mid Y_t = 0]$). In this setting we treat one of the two random variables as outside information, the likelihood under the null ignores this information while the alternative likelihood uses it:

$$E_t := \frac{\prod_{s=1}^{t} \left(\frac{n^t_{1|Y_s}}{n_{Y_s}}\right)^{X_t} \left(\frac{n^t_{0|Y_s}}{n_{Y_s}}\right)^{1-X_t}}{\left(\frac{n^t_1}{t}\right)^{n^t_1} \left(\frac{n^t_0}{t}\right)^{n^t_0}} \tag{20}$$

where $n_{1|Y_s}$ and $n_{0|Y_s}$ represent the number of times the variable X took on values of 1 and 0 respectively conditioned on the value of Y and $n_{Y_s} = n_{1|Y_s} + n_{0|Y_s}$. In the above e-process Y is assumed to be discrete. If it is not, then we can replace the empirical probabilities in the numerator with *any* estimator of $\mathbb{E}_{s-1}[X_s \mid Y_s]$.

The results of our hypothesis test suggest that there was a great deal of heterogeneity in the driver behavior during the initial months of the COVID-19 pandemic. Several of the same tags are found to be both significantly increasing and also decreasing in frequency. Some of these significant sequences of observations belonged to the same driver, as in Fig. 9. In these cases the sequence both increased and decreased at different points in time. However others, for example the residential building tag as shown in Fig. 10, were found to be increasing for some drivers *while simultaneously* decreasing for others. This suggests that the behavior changed differently across drivers. As before, this test does not explain *why* those differences exist: it could be that different drivers

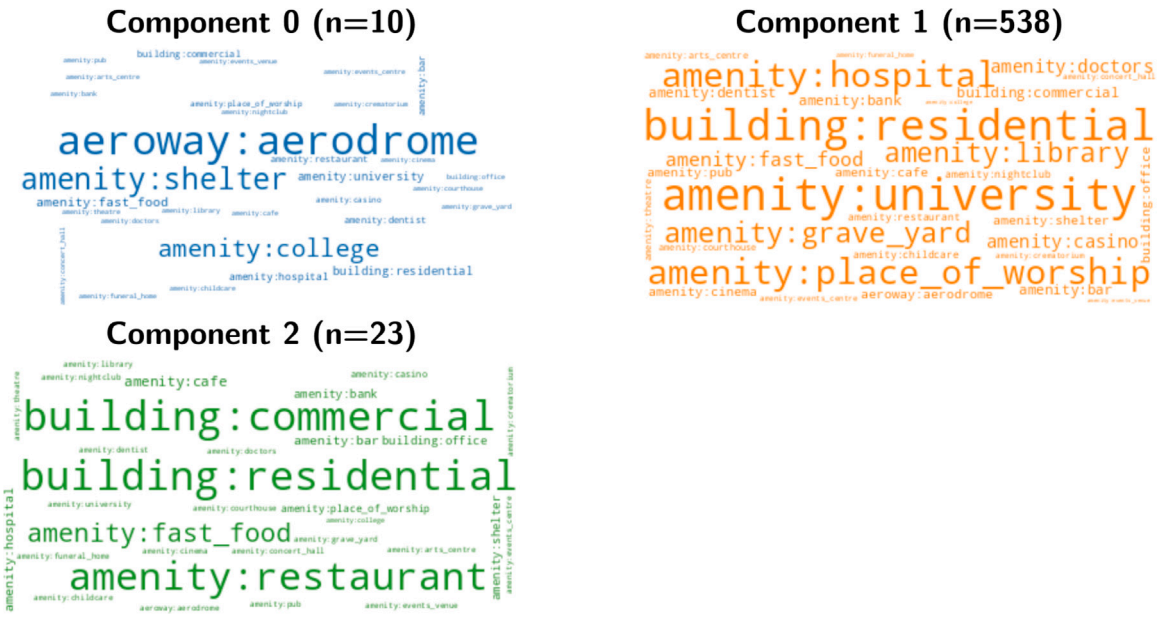

Fig. 11. Tag clouds for the 3 components extracted from the tag-seeking driver-tag e-values. The scale of each tag relates to its probability mass within the components. Titles are annotated with the number of drivers assigned to the cluster.

had different strategies or different responses to a changing environment or that a changing environment affected their different strategies in different ways, or even that they operate in different environments entirely.

We characterize the heterogeneity across drivers by fitting a Latent Dirichlet Allocation (LDA) model (Hoffman et al., 2010). LDA is often referred to as a topic model because it was developed to group a set of text documents into categories based on the words in the document. LDA simultaneously learns the categories or topics as a distribution over words (topic-word distribution), and the topic of the documents expressed as a distribution over topics (document-topic distribution). The documents are represented by a "document-term" matrix where the element at row i and column j is the number of times the word j appears in document i. In this setting we use LDA as a clustering method to identify behavioral profiles from the e-values of the drivers. We construct our document-term X as follows, for each driver i and tag j:

$$X_{i,j} = \ln\left(1 + \max_t E_t^{(i,j)}\right) \tag{21}$$

where $E_t^{(i,j)}$ is the value of our e-process for driver i and tag j at time t. We construct one document term matrix using the tag-seeking e-processes and another using the tag-avoiding e-processes and apply LDA to each separately. We take the maximum e-value to associate with each driver and tag the *largest* amount of evidence gathered in favor of the alternative over the observation period. Since e-values are multiplicative measures of evidence and the (pseudo-)count evidence that LDA considers are additive measures of evidence we apply a log transform to re-scale the e-values to conform with the assumptions of LDA. We set the document topic prior to 0.1 to encourage sparsity in the topic distribution assigned to each document, in our case this means that around 85% of drivers were assigned a topic distribution that placed at least 95% of its probability mass on a single topic. We assign component labels to each driver by taking the arg max of the document-topic distribution. The number of components (topics) was selected via cross-validation. We present the LDA results for tag-seeking behavior in Fig. 11.

The profiles extracted by LDA represent tags that are often sought together, but not necessarily at the same time. As shown in Fig. 11, Component 2, which counts restaurants, fast food, commercial buildings, and residences as its top tags appears to represent drivers who sought out food delivery. Component 0, on the other hand, which counts residences, the airport, commercial districts and colleges as its most likely tags, appears to represent drivers who sought passenger rides. Component 1 captures the majority of drivers and represents drivers who never accumulate much evidence against the null.

The profile analysis provided by LDA can be used to identify how individual driver trends are distributed over the entire set of drivers. This analysis is useful to the TNC operator as a way of segmenting drivers for intervention. For example, campaigns to boost goods delivery may be more successful with drivers who are expressing an increasing interest in food delivery (component 2). However, profile analysis is also useful for planners and regulators who are interested in driver behavior as collective rather than individual drivers. In this setting profile analysis reveals both the trends that are emerging and how widespread the trends appear to be. In fact, if we believe that the sample of drivers analyzed is representative of the driver population then cluster membership

provides an estimate of the population distribution over clusters. Understanding the distribution of the trend over drivers can help regulators determine the scope and scale of potential policy intervention.

Applying our procedure to the aggregated stops provides another system-level view which may be useful to TNC operators and public agencies as it directly represents the seeking and avoiding behavior of the drivers as a collective. In this setting our procedure identifies entities that are becoming more frequent at stops over time across the driver pool. Tag which are significant under the aggregate analysis have significant overlap with those identified in the components of the LDA analysis and those identified as significant in the individual driver analysis. The evidence, however, is stronger in the aggregate analysis as we are able to pool stop data across drivers.

That the evidence is stronger (i.e. the Bonferroni-corrected p-values are smaller) for the aggregated driver data is purely a function of the number of observations. It is not, importantly, a function of the time-scale of the observations. The unit of analysis for this approach is a stop and we are interested in how often a particular tag will be present at a stop. We receive an observation every time a driver makes a stop. It is important that these stops occur sequentially in time, but, apart from their order, it is not important how they are arranged in time. What this highlights is that this approach works on any time scale. What matters is how many observations are collected. This is why, even though the observations cover the exact same time window, our p-values for the aggregated stop sequences are so much smaller than the p-values for individual drivers: there are simply vastly more observations.

9. Applications

What is the utility of the information our approach is able to provide? And, to whom is it useful? First and foremost, drivers certainly had an interest in understanding what other drivers were doing when uncertainty was high *during* the early months of pandemic-related restrictions. This knowledge would have been useful information for them to have when deciding for themselves how (or whether) they should continue to engage with TNCs. Our method would have been useful to them because it provides detection of emerging behavioral trends, *at the time they are emerging*. The finding of significance is something that happens *while* the data is being collected, not after. In Figs. 7 and 8 and in Tables B.5 and B.6 the time points we identify as the moment at which the null was rejected are exactly the moments in time at which the significant result would have been available. It is neither necessary nor informative to wait for more data to be collected and so it is not a conclusion made in hindsight. This is a property of sequential hypothesis testing in general, but it is worth highlighting in our application. When we say, for example in Table B.5, we first detected that drivers were increasing their visits to fast food restaurants on April 21, 2020 it means that using *only* the data we actually had up to and including April 21, 2020, we were able to reject the null hypothesis. In other words, on April 21st of 2020 we would have been able to tell drivers that there was a statistically significant increase in fast food restaurant visits among drivers.

Perhaps the hallmark of the early pandemic response was a shared recognition that everything was changing but a shared uncertainty about what exactly was changing and how. TNCs knew the restrictions would fundamentally change their market, but may not have known how drivers and customers would respond. Our methodology would have been helpful for TNCs to determine *in real time* exactly what behavior drivers were changing and how. Early knowledge of emerging behavioral trends would have enabled them to validate their assumptions about how driver behavior would change with data and to make more informed decisions about how to invest in their platform: how to nudge drivers and customers to remain engaged with the platform. Individualized emerging behavioral trends would have also been useful for TNCs to determine possible individualized interventions. In Fig. 10 we noted an example of heterogeneous behavioral trends across drivers: some drivers were avoiding residential stops while others were seeking them out. In crafting strategies to engage drivers, it might have been useful for TNCs to detect for example, when an individual driver begins to shift away from or towards residential trips so that they may provide drivers with information about how to make the most of residential trips or, if they are avoiding these trips, other opportunities on the platform.

More broadly, one particularly useful application of this test is to detect when an event begins to occur more frequently than it has in the past, without the need to specify how frequently we should expect this event to occur nor the process by which this event is generated. When applied to TNC operations in particular, observers can detect with confidence that driver behavior has changed soon after it begins to change. The value of this information depends on who the observer is and what they are trying to do. We imagine that a data scientist at a TNC is trying to establish when, for example, an effort to increase the share of drivers involved in delivery (measured as, for example, the frequency with which a completed trip was delivery) has started to be effective. We imagine a regulator attempting to assess for example if and when a recent intervention at the curb is decreasing congestion or safety on the adjacent street. We imagine a data scientist at Gridwise who wants to detect when certain advice being offered to drivers is no longer as effective as it once was. Finally, we imagine that a TNC or government has implemented a policy under a set of assumptions about TNC driver behavior, perhaps for example, that drivers are more involved in rides than delivery. This methodology can be used to detect when these assumptions are no longer met.

In general we see two broad categories of use cases for this methodology. First, to establish with confidence and speed when some intervention begins to take effect in the desired way. Second, to establish with confidence and speed when an assumption that underlies a particular policy is no longer valid. Not only are both of these use cases powerful for regulators and organizations, but they are not methods to determine truth in hindsight, rather they are designed specifically to be monitored in real-time as data is being collected and to report with confidence as soon as the change is detected. Within choice modeling in particular, this methodology can be used to detect when the factors considered by a given choice model have a non-stationary relationship with the decisions observed over time.

In TNC operations, quickly detecting which kinds entities drivers are increasingly seeking and avoiding can help identify congestion problems and policy interventions both within government and within TNCs themselves. From the government perspective, this information can be used to assess TNC operations in the city without delving too deeply into operational metrics. Despite the fact that the data we use here neither was provided by a TNC nor contains any TNC-specific information, we are able to learn quite a bit about TNC driver behavior. When TNCs are reluctant to provide detailed operational data, governments may yet find ways to leverage our procedure to understand TNC operations from proxy data, as we have done here. As it concerns our particular work, policymakers and regulators can use this method to understand how TNC trends are changing in their cities, potentially in response to government regulation or changes in TNC policies or operations.

10. Conclusion

In this work we apply recent advances in sequential statistical hypothesis testing to extract behavioral changes in TNC driver behavior during the initial stages of the COVID-19 pandemic in the greater Pittsburgh area. We construct a composite null and composite non-parametric alternative sequential hypothesis test via an e-process capable of detecting when the probability of observing an event has increased or has decreased over a sequence of observations. Intuitively this e-process increases (accumulates evidence for the alternative) when the event starts occurring more frequently than it has in the past. The test is administered to the sequence of OpenStreetMap feature tags nearby service stops of 106 drivers over the first 9 months of the COVID-19 pandemic in Pittsburgh. The test found 112 driver tag pairs for which enough evidence existed to reject the null suggesting that the tag incidence probability was increasing (or decreasing) over time.

When a test finds that the sequence of tag incidences for a driver was becoming either significantly more or less frequent, it cannot tell us why. However, because the test places no assumptions on either the decision process in use by the driver or the environment, it can show us where to look for examples of changing behavior. Further analysis of these drivers could then yield insight into why certain changes were observed. Moreover, we find that certain tags became more frequent for some drivers while simultaneously becoming less frequent for others. This suggests heterogeneity in either or both of the driver decision processes or the environment that may be difficult to capture in a unified model. We describe this heterogeneity by fitting a Latent Dirichlet Allocation model to the driver-tag e-values to cluster drivers into behavioral trend profiles. We recover food delivery and passenger rides profiles consistent with our earlier analysis. Lastly, the e-process itself can be leveraged to test more specific hypotheses, ones that, in particular, attempt to address why certain tags occur more or less frequently for certain drivers.

In this work we restrict our focus to binary metrics. Although the e-process can be directly extended to accommodate any discrete metric, a key avenue of future work is to extend the testing methodology to bounded metrics. Processes that are martingales under our null that the sequence has constant (conditional) mean are simple to construct. In the language of Shafer and Vovk (2019) (see also Waudby-Smith and Ramdas, 2020) the challenge is to construct strategies to "bet against" the null so that under the alternative the e-process increases quickly.

A second key area of future work is a continued investigation into what was influencing the observed changes in behavior among these drivers. That the test does not explain the changes it identifies is on the one hand a strength of the methodology: because it does not rely on *any* assumptions on the behavior-generating process, failing to reject the null should be taken to mean that under *no model* did the *net* frequency of the event under study change significantly over time. However, from the practitioner's perspective it may be more interesting to focus on the drivers and tags for which we reject the null and for which the test offers no detailed explanation. In these cases the test identifies *what* needs to be explained and *when* the changes were first detected. Further investigations can thus be guided by these results.

CRediT authorship contribution statement

Matthew Battifarano: Literature review, Study conception and design, Data analytics, Programming, Analysis and interpretation of results, Manuscript preparation. **Sean Qian:** Study conception and design, Data acquisition, Analysis and interpretation of results, Manuscript preparation.

Acknowledgments

This work was supported by the National Science Foundation, United States grant CMMI-1931827. The authors would like to thank Gridwise Inc. for providing consultative resources for this research. Finally, the authors would like to extend thanks to Professors Glenn Shafer, Ruodu Wang, and Aaditya Ramdas for their course which introduced us to the statistical ideas we leverage in this paper and to Ojash Neopane for several helpful conversations.

Appendix A. Stop inference algorithm

We list the algorithm we developed to extract service stops from the GPS trajectory data in Algorithm 1.

Algorithm 1 Stop inference procedure

procedure StopInference($\{(v_i, t_i, d_i)\}_{i=1}^{m}$)
  $n \leftarrow 0$ ▷ Initialize the service stop counter
  for $i \in \{1, \cdots, m\}$ **do**
    $S_i \leftarrow \arg\max_s \Pr[S_i = s \mid V = v_i]$ ▷ Infer S_i
    if $i > 1$ **then**
      $C_{i-1} \leftarrow \arg\max_c \Pr[C_{i-1} = c \mid T = t_{i-1}, D = d_{i-1}, S_{i-1}S_i = 1]$ ▷ Infer C_i
      if $C_i = 0$ **then**
        $n \leftarrow n + 1$ ▷ Increment stop counter
      end if
    end if
    if $S_i = 1$ **then** ▷ Assign a service stop label
      $N_i \leftarrow n$
    else
      $N_i \leftarrow -1$
    end if
  end for
  return $\{N_i\}_{i=1}^{n}$
end procedure

Table B.4
Summary of test results.

Tag	Seeking		Avoiding	
	Best p-value	n	Best p-value	n
aeroway:aerodrome	**3.95e−62**	**10**	**8.94e−66**	**17**
amenity:arts_centre	1.48e+00		6.91e−01	
amenity:bank	1.30e−03		6.43e−04	
amenity:bar	2.53e−05		3.00e−04	
amenity:cafe	1.17e−01		**1.64e−08**	**1**
amenity:casino	**5.31e−10**	**1**	**7.10e−12**	**1**
amenity:childcare	1.85e+00		1.80e−01	
amenity:cinema	3.57e−04		5.54e−01	
amenity:college	**1.62e−64**	**1**	1.41e−02	
amenity:concert_hall	2.00e+00		2.00e+00	
amenity:courthouse	1.25e+00		5.16e−04	
amenity:crematorium	2.00e+00		2.00e+00	
amenity:dentist	**2.25e−06**	**1**	6.34e−02	
amenity:doctors	**1.40e−06**	**1**	2.08e−04	
amenity:events_centre	2.00e+00		2.00e+00	
amenity:events_venue	2.00e+00		1.27e−02	
amenity:fast_food	**2.46e−11**	**4**	**4.37e−18**	**7**
amenity:funeral_home	2.00e+00		2.00e+00	
amenity:grave_yard	**1.54e−15**	**1**	5.11e−01	
amenity:hospital	8.64e−04		**2.46e−06**	**1**
amenity:library	**8.63e−10**	**1**	5.96e−01	
amenity:nightclub	1.77e+00		1.76e−01	
amenity:place_of_worship	**5.20e−19**	**1**	9.03e−04	
amenity:pub	3.04e−01		3.42e−04	
amenity:restaurant	**2.87e−161**	**4**	**1.25e−20**	**4**
amenity:shelter	**1.26e−41**	**5**	**1.97e−18**	**3**
amenity:theatre	1.25e+00		2.97e−02	
amenity:university	**1.30e−08**	**1**	**5.22e−14**	**1**
building:commercial	**2.57e−130**	**8**	**5.13e−36**	**5**
building:residential	**2.47e−41**	**17**	**3.39e−46**	**15**
building:office	1.25e+00		**2.15e−06**	**1**

Appendix B. Extended summary of test results

Each tested tag is listed in Table B.4. For each tested tag, we report the smallest p value (before Bonferoni correction) over the set of drivers tested as well as the number of drivers for which the tag was found to be significant for each of the two sides of the two-sided test, "seeking" and "avoiding" respectively.

Tables B.5 and B.6 list the tags found to be significant, their Bonferroni-corrected p value, as well as when significance was first detected, by the procedure applied to the aggregated sequence of stops.

Table B.5
Bonferroni-corrected p-values for significant tag-seeking behavior.

Tag	p value	First detected
amenity:fast_food	3.43e−104	2020-04-21 19:46:07.441
amenity:restaurant	1.22e−76	2020-05-05 18:40:57.001
amenity:college	1.58e−69	2020-09-09 14:57:19.000
building:commercial	2.05e−60	2020-04-14 02:30:59.047
amenity:grave_yard	9.46e−39	2020-07-12 22:41:35.114
amenity:shelter	1.24e−29	2020-05-30 19:49:09.731
aeroway:aerodrome	3.21e−29	2020-09-10 22:53:09.000
amenity:university	1.46e−18	2020-09-10 07:16:22.163
amenity:casino	1.33e−17	2020-08-08 18:34:05.000
building:residential	3.92e−16	2020-05-01 17:06:29.999
amenity:bar	2.93e−12	2020-10-31 18:52:31.048
amenity:cafe	6.54e−11	2020-05-23 12:52:53.143
amenity:hospital	1.41e−10	2020-09-18 07:56:58.000
building:office	1.81e−07	2020-11-02 15:11:41.000
amenity:theatre	5.94e−07	2020-06-08 20:47:15.070

Table B.6
Bonferroni-corrected p-values for significant tag-avoiding behavior.

Tag	p value	First detected
aeroway:aerodrome	0.00e+00	2020-03-27 21:25:11.512
amenity:casino	2.63e−31	2020-03-22 20:09:48.494
amenity:dentist	3.30e−26	2020-05-26 09:22:41.397
building:residential	4.02e−24	2020-08-12 19:34:19.999
amenity:arts_centre	1.71e−15	2020-06-11 09:51:00.001
amenity:university	1.92e−08	2020-06-26 23:11:47.998

Appendix C. Increasing probability as the null hypothesis

In this work we treat the hypothesis that the sequence is i.i.d. Bernoulli with unknown mean as the null hypothesis and the hypothesis that the mean is increasing (decreasing) as the alternative. We choose this assignment of null and alternative for two reasons one conceptual and one technical. Conceptually, the null should be chosen to represent the uninteresting case: the intervention has *no effect*, there is *no relationship* between two measurements, or, in the case of the very first null hypothesis, Muriel Bristol is no better than chance at distinguishing which of tea or milk was first added to a cup (Fisher, 1936). In our setting we argue that the uninteresting case is when the probability of a service stop being nearby an entity with a given tag is constant in time. The technical reason is that in order for our statistic to be an e-process we must be able to compute the supremum in the denominator Eq. (14). This may be done in closed form when the null is i.i.d. Bernoulli, but when the null is instead the set of Bernoulli sequences with non-decreasing mean we may use the following convex optimization to identify the maximum likelihood under the null;

$$\ell_+^*(X) = \max_{\mathbf{p}\in\mathbb{R}^n} \quad \sum_{i=1}^{n} X_i \log(p_i) + (1 - X_i)\log(1 - p_i) \tag{C.1a}$$

$$\text{s.t.} \quad 0 \le p_1 \le p_2 \le \cdots \le p_n \le 1 \tag{C.1b}$$

An analogous program may be used for the hypothesis that the probabilities are non-increasing;

$$\ell_-^*(X) = \max_{\mathbf{p}\in\mathbb{R}^n} \quad \sum_{i=1}^{n} X_i \log(p_i) + (1 - X_i)\log(1 - p_i) \tag{C.2a}$$

$$\text{s.t.} \quad 1 \ge p_1 \ge p_2 \ge \cdots \ge p_n \ge 0 \tag{C.2b}$$

It should be noted that the i.i.d. Bernoulli product distribution is a member of both of the above constraints sets. As a result, i.i.d. Bernoulli makes a poor alternative hypothesis when either the non-decreasing or non-increasing hypothesis is used as the null. Instead, the non-decreasing and non-increasing hypotheses can be used interchangeably as null and alternative to one another, or the non-stationary alternative $p^{\pm}$ defined in Eq. (15) may be used.

References

Agrawal, S., Goyal, N., 2012. Analysis of Thompson sampling for the multi-armed bandit problem. In: Conference on Learning Theory. In: JMLR Workshop and Conference Proceedings, pp. 39–1.

Alemi, F., Circella, G., Handy, S., Mokhtarian, P., 2018. What influences travelers to use Uber? Exploring the factors affecting the adoption of on-demand ride services in California. Travel Behav. Soc. 13, 88–104.
Allegheny County Airport Authority, 2020. Pittsburgh international airport summary of airline traffic april 2020. URL https://flypittsburgh.com/acaa-corporate/about/airport-statistics/.
Arora, S., Doshi, P., 2021. A survey of inverse reinforcement learning: Challenges, methods and progress. Artificial Intelligence 297, 103500.
Bansal, P., Sinha, A., Dua, R., Daziano, R.A., 2020. Eliciting preferences of TNC users and drivers: evidence from the United States. Travel Behav. Soc. 20, 225–236.
Ben-Akiva, M.E., 1985. Discrete Choice Analysis Theory and Application to Travel Demand. MIT Press Series in Transportation Studies, vol. 9, MIT Press, Cambridge, Mass.
Berliner, R., Tal, G., 2018. What drives your drivers: An in-depth look at Lyft and Uber drivers. UC Davis Institute of Transportation Studies. https://steps.ucdavis.edu/wpcontent/uploads/2018/02/BERLINER-TAL-What-Drives-Your-Drivers.pdf.
Bhat, C.R., 1997. Recent methodological advances relevant to activity and travel behavior analysis. In: International Association of Travel Behavior Research Conference. Austin, Texas, Citeseer.
Cohn, D.A., Ghahramani, Z., Jordan, M.I., 1996. Active learning with statistical models. J. Artificial Intelligence Res. 4, 129–145.
de Ruijter, A., Cats, O., Kucharski, R., van Lint, H., 2022. Evolution of labour supply in ridesourcing. Transportmetrica B: Transp. Dyn. 10 (1), 599–626.
Dias, F.F., Lavieri, P.S., Garikapati, V.M., Astroza, S., Pendyala, R.M., Bhat, C.R., 2017. A behavioral choice model of the use of car-sharing and ride-sourcing services. Transportation 44 (6), 1307–1323.
Fisher, R.A., 1936. Design of experiments. Br. Med. J. 1 (3923), 554.
Greene, W., 2009. Discrete choice modeling. In: Palgrave Handbook of Econometrics. Springer, pp. 473–556.
Grünwald, P., de Heide, R., Koolen, W.M., 2020. Safe testing. In: 2020 Information Theory and Applications Workshop. ITA, IEEE, pp. 1–54.
Hall, J.V., Krueger, A.B., 2018. An analysis of the labor market for Uber's driver-partners in the United States. Ilr Rev. 71 (3), 705–732.
Head, M.L., Holman, L., Lanfear, R., Kahn, A.T., Jennions, M.D., 2015. The extent and consequences of p-hacking in science. PLoS Biol. 13 (3), e1002106.
Hoffman, M., Bach, F., Blei, D., 2010. Online learning for latent dirichlet allocation. Adv. Neural Inf. Process. Syst. 23.
Howard, S.R., Ramdas, A., 2019. Sequential estimation of quantiles with applications to A/B-testing and best-arm identification. arXiv preprint arXiv:1906.09712.
Howard, S.R., Ramdas, A., McAuliffe, J., Sekhon, J., 2020. Time-uniform Chernoff bounds via nonnegative supermartingales. Probab. Surv. 17, 257–317.
Keane, M.P., Wolpin, K.I., 2009. Empirical applications of discrete choice dynamic programming models. Rev. Econ. Dyn. 12 (1), 1–22. http://dx.doi.org/10.1016/j.red.2008.07.001, URL https://www.sciencedirect.com/science/article/pii/S1094202508000318.
Kim, J., Kim, J.H., Lee, G., 2022. GPS data-based mobility mode inference model using long-term recurrent convolutional networks. Transp. Res. C 135, 103523.
Lavieri, P.S., Bhat, C.R., 2019. Investigating objective and subjective factors influencing the adoption, frequency, and characteristics of ride-hailing trips. Transp. Res. C 105, 100–125.
McDonald's Corporation, 2020. Form 10-Q. URL https://www.sec.gov/ix?doc=/Archives/edgar/data/0000063908/000006390820000063/mcd-6302020x10q.htm.
McFadden, D., 1974. The measurement of urban travel demand. J. Public Econ. 3 (4), 303–328.
McFadden, D., et al., 1973. Conditional Logit Analysis of Qualitative Choice Behavior. Institute of Urban and Regional Development, University of California Oakland.
Miller, J., Nie, Y., Stathopoulos, A., 2017. Crowdsourced urban package delivery: Modeling traveler willingness to work as crowdshippers. Transp. Res. Rec. 2610 (1), 67–75.
OpenStreetMap contributors, 2022. OpenStreetMaps Data. https://www.openstreetmap.org.
Punel, A., Stathopoulos, A., 2017. Modeling the acceptability of crowdsourced goods deliveries: Role of context and experience effects. Transp. Res. E 105, 18–38.
Ramdas, A., Ruf, J., Larsson, M., Koolen, W., 2020. Admissible anytime-valid sequential inference must rely on nonnegative martingales. arXiv preprint arXiv:2009.03167.
Ramdas, A., Ruf, J., Larsson, M., Koolen, W.M., 2022. Testing exchangeability: Fork-convexity, supermartingales and e-processes. Internat. J. Approx. Reason. 141, 83–109.
Servizi, V., Petersen, N.C., Pereira, F.C., Nielsen, O.A., 2020. Stop detection for smartphone-based travel surveys using geo-spatial context and artificial neural networks. Transp. Res. C 121, 102834.
Shafer, G., 2019. The language of betting as a strategy for statistical and scientific communication. arXiv preprint arXiv:1903.06991.
Shafer, G., Shen, A., Vereshchagin, N., Vovk, V., 2011. Test martingales, Bayes factors and p-values. Statist. Sci. 26 (1), 84–101.
Shafer, G., Vovk, V., 2019. Game-Theoretic Foundations for Probability and Finance, Vol. 455. John Wiley & Sons.
Shaffer, J.P., 1995. Multiple hypothesis testing. Annu. Rev. Psychol. 46 (1), 561–584.
Shin, J., Ramdas, A., Rinaldo, A., 2022. E-detectors: a nonparametric framework for online changepoint detection. http://dx.doi.org/10.48550/ARXIV.2203.03532, URL https://arxiv.org/abs/2203.03532.
Smith, R.C., 2013. Uncertainty Quantification: Theory, Implementation, and Applications, Vol. 12. Siam.
Sutton, R.S., Barto, A.G., et al., 1998. Introduction to Reinforcement Learning. MIT press Cambridge.
Uber Technologies, 2019. Form 10-K. URL https://www.sec.gov/ix?doc=/Archives/edgar/data/0001543151/000154315120000010/fy2019q410kfinancialst.htm.
Uber Technologies, 2020. Form 10-Q. URL https://www.sec.gov/ix?doc=/Archives/edgar/data/0001543151/000154315120000029/uber-20200630.htm.
Ville, J., 1939. Etude critique de la notion de collectif. Gauthier-Villars, Paris.
Vovk, V., Wang, R., 2021. E-values: Calibration, combination and applications. Ann. Statist. 49 (3), 1736–1754.
Wald, A., 1945. Sequential tests of statistical hypotheses. Annu. Math. Stat.
Wasserman, L., Ramdas, A., Balakrishnan, S., 2020. Universal inference. Proc. Natl. Acad. Sci. 117 (29), 16880–16890.
Waudby-Smith, I., Ramdas, A., 2020. Estimating means of bounded random variables by betting. arXiv preprint arXiv:2010.09686.
Xiao, G., Juan, Z., Zhang, C., 2016. Detecting trip purposes from smartphone-based travel surveys with artificial neural networks and particle swarm optimization. Transp. Res. C 71, 447–463. http://dx.doi.org/10.1016/j.trc.2016.08.008, URL https://www.sciencedirect.com/science/article/pii/S0968090X16301425.
Ziebart, B.D., Maas, A.L., Bagnell, J.A., Dey, A.K., et al., 2008. Maximum entropy inverse reinforcement learning. In: Aaai. Chicago, IL, USA, 8, pp. 1433–1438.