DESCRIPTION AND EXPLANATION

PAUL DELATTE

University of Southern California, *delatte@usc.edu*

In this paper, we formalize the existence of two different tasks in model-based inferential statistics: description and explanation. While both tasks lead to sensibly similar statistical procedures, they rely on fundamentally different modeling approaches. We map the task of explanation to what are commonly known as inverse problems in statistics. This allows us to clarify the nature, structure, and interpretation of these problems where the objective is to recover some indirectly observable object whose existence derives from a preliminary stochastic model linking observable to unobservable processes. By contrast, we delineate a large class of problems where inversion plays no role. The reason is that they tackle the simpler task of description that only consists in assigning a probability measure to its realizations. This descriptive task can be performed without committing to an explanatory model nor to the existence of unobservable processes.

KEYWORDS: statistical models; forward problems; inverse problems; identification; misspecification.

1 Introduction

1.1 Problem statement and solution overview

A standard way to introduce model-based inferential problems in statistics is through parametrized statistical models of the form $\{P^{\theta} : \theta \in \Theta\}$ where P^{θ} is some probability measure and Θ some parameter set. The objective is then set to recover the true unknown parameter $\theta_0 \in \Theta$ based on a finite set values that are assumed to be the realizations of some random variables with joint distribution P^{θ_0}. This starting point has the benefit of brevity and mathematical elegance: it can fit many inferential problems and can be used to obtain widely applicable and easily comparable theoretical results. Despite the familiarity of statisticians and econometricians with such a starting point, parametrized statistical models as indexed families of probability measures are not easy objects to handle nor to interpret[1].

To illustrate this point, consider one of the most elementary statistical models. For a single observation z_1 of a random variable Z_1, we suppose that the distribution of Z_1 lies in

$$\{N(\theta, 1) : \theta \in \mathbb{R}\}.$$

This version: February 20, 2025.

[1] It should be remembered that an indexed family of probability measures $\{P^{\theta} : \theta \in \Theta\}$ implies the existence of an unindexed family of measures $\mathscr{P}$, an index set Θ, and a surjective function $I : \Theta \to \mathscr{P}$ called index map.

This defines a statistical model for Z_1 as well as a statistical problem. The problem consists in gaining knowledge of the unknown true parameter $\theta_0 \in \mathbb{R}$ based on z_1 where we suppose that z_1 is the realization of Z_1 such that

$$Z_1 \sim N(\theta_0, 1). \tag{1.1}$$

While this problem is well-defined mathematically, a number of preliminary questions are usually left unanswered: How is one to interpret the choice of the possible distributions, the choice of the parametrization, the choice of the index set? What is the objective of the modeler by framing such a problem? How should the resulting statistical claims (confidence intervals or hypothesis tests) be interpreted? To add to the confusion, note that the simple statistical model considered above is often obtained from first considering the following equation

$$Z_1 = \theta_0 + X_1 \tag{1.2}$$

where X_1 is a random variable assumed distributed according to $N(0, 1)$ and θ_0 is a constant random variable whose value is left unknown up to the hypothesis that it lies in $\mathbb{R}$. The resulting statistical problem is the same as the one obtained from the previously considered statistical model: gaining knowledge of θ_0 based on z_1 which is assumed to be a realization of Z_1 with distribution $N(\theta_0, 1)$ as given by the hypothesized equation. For this reason, the statistical model $\{N(\theta, 1) : \theta \in \mathbb{R}\}$ and the equation $Z_1 = \theta_0 + X_1$ with $X_1 \sim N(0, 1)$ are often considered and used interchangeably. This raises a number of additional questions that are more troubling than the first ones: Are the two structures identical? If not, what is the objective of the model defined by the equation above? What is then the interpretation of the statistical results obtained from it? Naturally, the model defined by the equation above is not the only one that generates the statistical problem under consideration. Infinitely many other equations of varying similarity lead to the same problem: e.g., for any $K > 1$, $Z_1 = \theta_0 + \sum_{k=1}^{K} X_k$ where $X_1, X_2, \ldots, X_K$ are assumed mutually independent and distributed according to $N(0, 1)$. This only reinforces the initial puzzle: How to interpret all these different models that lead to the same statistical decisions?

It is the objective of this paper to answer this question. By doing so, we shed light on the existence of two different problems within model-based inferential statistics that logically precede any parametrized statistical model. The distinction we establish helps clarify a number of intricate issues in inferential statistics such as identification, specification, and model validity. Before dwelling into formal definitions, we briefly introduce the two inferential problems that are the object of this paper. To do so, we start by going back to the fundamental building blocks of any model-based inferential problem. We have:

1. a finite set of observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$;

2. a modeling assumption that these observations are the realizations of some random variables $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ with some unknown joint distribution $P_{\mathbf{Z}}$.

From these building blocks, we claim that it is possible to consider two distinct tasks.

A. **Description**. The objective is to find a plausible distribution for the random phenomenon from which we obtained the observations $z_1, z_2, \ldots, z_n$. The descriptive nature of the task follows from the simple act of associating a distribution to its realizations.

B. **Explanation**. The object of interest is not the random phenomenon from which we obtained the observations $z_1, z_2, \ldots, z_n$, but a theoretical term that depends only partially on them in virtue of an a priori stochastic model linking $\mathbf{Z}$ to some unobservable processes. The explicative nature of the task follows from hypothesizing an unobservable structure from which a theoretical object is defined and made recoverable from the realizations of an adjacent observable process.

The goal of the next two sections is to map these two tasks to properly delineated problems. This will provide a direct answer to the questions that have been previously raised, not only for mean estimation in the Gaussian family with known variance but for most model-based inferential problems in statistics. For this purpose, we introduce two simple structures that generate most parametrized statistical models in practice and properly formalize the two tasks. An interesting feature of our analysis is that the structure sustaining explanation is seen to generate what are already known as inverse problems in statistics. Since these problems have long evaded any proper formalization in spite of their high prevalence in applications, it is another contribution of our paper to make clear the nature, structure, and interpretation of these problems. Moreover, while the task of description could start from parametrized statistical models, we still believe that there exists a more fundamental structure that stands behind them. The structure we exhibit is as simple as possible and directly captures how statistical models for observable processes are formed in practice. One may deem the exercise superfluous in this case, but we believe that it helps clarify a number of issues in inferential statistics that confuse theoreticians and practitioners alike.

Remark 1.1. The focus of this paper is model-based (or sampling-based) inferential statistics where the observations are assumed to be realizations of some random variables (called population or super-population) of which some features are to be statistically recovered. This is the most standard approach to inferential statistics to the point where the qualifier "model-based" is often dropped. Model-based inference is to be distinguished from design-based (or randomization-based) inference which is the favored solution whenever one has direct control over the sampling. In this case, it is possible to perform inference without committing to a sampling (super-)population by only leveraging the randomness of the sampling mechanism. This typically happens in the study of surveys and experiments (see footnote in Example 3.7).

Remark 1.2. Another terminological clarification is in order in view of a conflicting use of the adjective "descriptive" in statistics. The forward problems we consider are inferential problems: they lead to an inductive step and call for some quantification of uncertainty around it. They should not be viewed as part of "descriptive statistics" where there is no attempt at induction nor at the quantification of uncertainty around inductive claims. We believe it to be uncontroversial that, in spite of the mild modeling assumption of a theoretical population inherent to model-based inferential problems, some of these problems are descriptive in nature. This descriptive task is understood as assigning a probability measure to some observable phenomena (and should be compared to the explicative task of inverse problems where a model explaining how such a probability measure is obtained from unobservable random processes is hypothesized).

Remark 1.3. The two structures we exhibit should not alter standard inferential practices in so far as these structures logically precede parametrized statistical models (see Remark 2.3 and Remark 3.4). In particular, our analysis is completely agnostic to the inferential engine used (Fisherian, frequentist, Bayesian) and to the nature of the statistical claims considered (hypothesis tests and p-values, point estimates and confidence intervals, posterior probabilities and credible intervals). However, we expect the structures we exhibit to alter modeling practices. In particular, they should help make clear that the validity and interpretation of statistical claims in model-based inferential problems are conditional on the correct specification of a structure that is not reducible to a parametrized statistical model. Moreover, the conditioning sets under which these statistical claims hold true differ significantly depending on the structure considered (see Remark 2.5 and Remark 3.6) and thus on the objective of the modeler, whether the goal is a description or the recovery of a theoretical object from an explanation.

1.2 Literature review

The use of indexed family of probability measures as the traditional starting point for model-based inferential problems in statistics can be easily perceived by skimming through most statistical textbooks independently of the era, school, and inferential paradigm – see, e.g., Ferguson (1967), Berger (1985), Lehmann and Casella (1998), Lehmann and Romano (2005), van der Vaart (2000), Robert (2007), Shao (2008), and Efron and Hastie (2021) for a small fairly representative sample.

On this universal backdrop, models defined through functional equations of the type (1.2) are often introduced but are almost always taken as equivalent to the parametrized statistical model they indirectly lead to (in virtue of their simple bijective parametrizations). It is only for more complex problems that proper attention has been paid to the difference, due to sui generis complications such as identification issues. This has been accompanied by the development of the notion of inverse problems in statistics (see Cavalier (2011), Florens and Simoni (2017)) where some specific features of these problems have been clearly exhibited. However, in the rare cases the notion has been explicitly formalized, it has always been inextricably tied to inverse problems in the sciences where an operator needs to be inverted: that is, find $f \in F$ such that $A(f) = g$ where A is a known operator and g is known up to some stochastic disturbance. This is naturally the same approach that has been used by physical scientists when connecting physical inverse problems to statistical problems (see Evans and Stark (2002), Tarantola (2005), Aster, Borchers, and Thurber (2018) going as far back as Sudakovand and Khalfin (1964)). This, however, does not give justice to the diversity of inverse problems considered in statistics that are, in the words of Cavalier (2011), all "[t]hese are problems where we have indirect observations of an object [...] that we want to reconstruct". We argue in this paper that there is a way to directly formalize the notion without considering as a starting point a deterministic inverse problem. This allows to fit any stochastic model of the type (1.2) within the frame of statistical inverse problems, independently of its connection or not to a deterministic inverse problem. This makes possible the consideration of problems where the measurement issue is not due some stochastic disturbance (as in physical problems) but to theoretical reasons (as in social science problems as well as in causal inference problems). The defining structure for inverse problems in statistics we exhibit is relatively simple and depends mostly on the formulation of a stochastic model in terms of unobservable processes. It allows, however, to unify many practices and clarify further many features specific to these problems. It finally makes clear that the name of inverse problems in statistics does not come from inverting an operator but from "undoing" the effects of unobservable processes (as clearly perceived in Giné and Nickl (2021) pp. 3-4).

The general formulation of inverse problems through stochastic models then begs the question: Are all model-based inferential problems inverse? If not, what are the other problems about? It is hard not to answer an emphatic "no" to the first question by even considering the simplest inferential problem – non-parametric mean estimation. To make sense of this negative answer, we exhibit a very simple structure that elegantly capture a large class of model-based inferential problems where inversion plays no role. This follows from and confirm the descriptive nature of such problems. As far as we know, the forward structure we introduce, in spite of its simplicity, has never been considered so directly, especially in light of the inverse structure of inverse problems that is made explicit in this paper. A different attempt at formalizing similar ideas can be found in Evans and Stark (2002) but the authors only ever define forward problems in statistics in terms of parametrized statistical models (hence begging the question as to the origin of such models). The forward structure in this paper makes clear the logical precedence of measures as complete descriptions before any attempt of parametrization. Since nothing is new under the sun, the differences we highlight may be deem too obvious – if so, we believe the objective of the paper has been at least partly accomplished.

More generally, we grapple in this paper with the existence of different tasks within statistics. Naturally, we are not the first one to face the issue. The idea of a fundamental difference between inferential problems lays latent in the diversity of statistical applications. Several papers have tried to explicitly characterize this difference and offer support for one or another position. The most influential contribution in this direction is without contest the "two cultures" paper of Leo Breiman (Breiman (2001)). The author identifies two tasks within statistics, "prediction" and "information [extraction]", and two criteria to judge inferential results, "prediction accuracy" and "interpretability", and then compare the "data modeling culture" and the "algorithmic modeling culture" to serve these objectives. Breiman, however, takes stochastic models as a given and remains mostly silent about the meaning of "information [extraction]" and "interpretability". Our paper can be understood as showing that "information [extraction]" takes at least two distinct forms, description and explanation, and that the difference between these two tasks can be formalized through the difference between forward problems and inverse problems. As far as we know, we are the first one to confront so directly the question of the nature of "information [extraction]" within statistics. This discussion naturally fits in a larger epistemological literature where the distinction has been amply considered (see, e.g., Woodward and Ross (2021) and the references therein).

2 Description as forward problems

2.1 Definition and examples

In model-based inferential statistics, one of the objectives is to simply provide a description of some observable phenomenon. This translates in finding the right predicate, that is, the right distribution from which the observations of the phenomenon are supposed to be drawn. This is what we define as a forward problem in statistics. Given a set of observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ assumed to be the realizations of random variables $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$, the object of interest is their unknown joint distribution P_Z which captures the complete description of the phenomenon under study. To make the problem tractable, it is natural to restrict the possible descriptions as well as the quantity of information within each description. A forward problem in statistics hence consists in choosing successively:

1. a set of probability measures $\mathscr{P}$ that is supposed to contain $P_{\mathbf{Z}}$: this delineates the range of possible descriptions for the phenomenon under study;

2. a well-defined function $f : \mathscr{P} \to E$ where E is an arbitrary set: this restricts the quantity and type of information that should be considered for any description.

The statistical problem then consists in gaining knowledge of $f(P_{\mathbf{Z}})$ from the observations. It is important to note that the object of interest remains $P_{\mathbf{Z}}$, which captures the complete description of the phenomenon under study. In particular, if we were able to know $P_{\mathbf{Z}}$, we could directly compute any well-defined quantity $f(P_{\mathbf{Z}})$ from it. However, because the problem of recovering $P_{\mathbf{Z}}$ may be untractable, it is common to consider a sub-problem of lesser difficulty. The function f can thus be understood as a filter to make the statistical problem simpler and the image $f(P_{\mathbf{Z}})$ as providing a partial description of the phenomenon. The choice of $\mathscr{P}$ and f can be seen as trading off the difficulty of the statistical problem with the completeness of the intended description. For some choice of $\mathscr{P}$, it sometimes happens that $f(P_{\mathbf{Z}})$ completely characterize $P_{\mathbf{Z}}$, in which case recovering $P_{\mathbf{Z}}$ and recovering $f(P_{\mathbf{Z}})$ are one and the same problem.

Definition 2.1 (Forward Problem)**.** Consider a set of observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ assumed to be the realizations of a random process $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ with unknown joint distribution $P_{\mathbf{Z}}$. A forward problem for $P_{\mathbf{Z}}$ is characterized by a pair $(\mathscr{P}, f)$ where $\mathscr{P}$ is a set of probability measures that is supposed to satisfy $P_{\mathbf{Z}} \in \mathscr{P}$ and $f : \mathscr{P} \to E$ is a function with E an arbitrary set. The statistical problem consists in gaining knowledge of $f(P_{\mathbf{Z}})$ using $\mathbf{z} = (z_1, z_2, \ldots, z_n)$.

Beyond the heaviness of the definition, the procedure is the most straightforward one could consider in model-based inferential statistics. It is legitimate to wonder why such an intuitive procedure need to be formalized. We claim, however, that doing so helps make clear the nature of many problems in statistics, their hypotheses, and what is obtained from them. To illustrate this, we exhibit the forward structure of four standard inferential problems: non-parametric mean estimation, normal mean estimation, non-parametric regression, and linear regression. For each of these problems, we make clear the choice of $\mathscr{P}$ and f as well as the intended description motivating them.

Remark 2.1. When the random variables $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$, from which the observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ are realizations, are further assumed to be independently and identically distributed (i.i.d.) according to a random variable Z with distribution P_Z, we should abuse notation and write the forward problem $(\mathscr{P}, f)$ directly for P_Z instead of $P_{\mathbf{Z}}$. The original forward problem can then be obtained by expanding $\mathscr{P}$ through the product measure operator and by defining f pointwise on the Cartesian product. While it is natural to abuse notation this way, it is important to remember the restriction when considering the plausibility of the resulting descriptions.

Example 2.2 (Non-parametric Mean Estimation)**.** The problem of non-parametric mean estimation consists in estimating the mean $\mathbb{E}[X] < \infty$ of a real-valued random variable X with distribution P_X based on the realizations $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ of an i.i.d. sample $(X_1, X_2, \ldots, X_n)$ that is assumed distributed according to P_X. This is a standard inferential problem with an explicit forward structure. Due to the i.i.d. assumption, we directly state the forward structure $(\mathscr{P}, f)$ for P_X according to Remark 2.1. The possible descriptions of the phenomenon under study are restricted by selecting

$$\mathscr{P} = \left\{P \in \mathcal{M}(\mathbb{R}) : \int_{\mathbb{R}} x \, dP(x) < \infty\right\}.$$

The restriction of the possible descriptions comes from a mild non-parametric assumption (integrability) and an i.i.d. assumption. Once the set $\mathscr{P}$ is chosen, the quantity and type of information to extract are restricted by considering the mean operator, that is,

$$f \colon \mathscr{P} \to \mathbb{R},\ P \mapsto \int_{\mathbb{R}} x \, dP(x).$$

The expected value exists and is unique, so the function f is well-defined and hence the forward problem is also well-defined. The statistical problem then consists in estimating $\int_{\mathbb{R}} x \, dP_X(x) = \mathbb{E}[X]$ using the sample $(x_1, x_2, \ldots, x_n)$ assumed to be independent realizations of $P_X \in \mathscr{P}$. The choice of f is typically motivated by the fact that inference for the full distribution P_X (an infinite-dimensional object) may be too hard a problem given the data at hand, while the expected value (a one-dimensional object) may be more tractable and provide sufficient information. The function f is not bijective in this case: the description obtained is only partial. Moreover, the statistical problem remains difficult due to the non-parametric assumption in $\mathscr{P}$: for instance, Bahadur and Savage (1956) showed that, due to the dependence of the mean on the tails, the non-parametric assumption prevents the existence of "effective" statistical procedures for $\mathbb{E}[X]$.

Example 2.3 (Mean Estimation in a Gaussian Family)**.** The problem of mean estimation in a Gaussian family (with known variance) is a sub-problem directly obtained from non-parametric mean estimation. The problem consists in estimating $\mathbb{E}[X] = \mu_0$ given observations $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$

assumed to be independent realizations of X with distribution $N(\mu_0, 1)$. This problem has an explicit forward structure $(\mathscr{P}_G, f)$ where the set of possible descriptions $\mathscr{P}_G$ is seen to be a strict subset of the set $\mathscr{P}$ of possible descriptions for non-parametric mean estimation, namely

$$\mathscr{P}_G = \left\{N(\mu, 1) : \mu \in \mathbb{R}\right\} \subset \mathscr{P}.$$

The information filter is, however, unchanged and consists in the restriction of the mean operator to $\mathscr{P}_G$, namely

$$f\colon \mathscr{P}_G \to \mathbb{R},\ P \mapsto \int_{\mathbb{R}} x\, dP(x).$$

Since the mean also exists and unique for this subset, the function f and the forward problem are well-defined. The statistical then consists in estimating $\mathbb{E}[X] = \mu_0$ based on observations $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ assumed to be independent realizations of $X \sim N(\mu_0, 1)$. This problem is much simpler in virtue of the parametric restriction $\mathscr{P}_G$: the Gaussian assumption ensures control of the tails preventing the behaviors in Bahadur and Savage (1956) to manifest. Moreover, in virtue of the restriction to $\mathscr{P}_G$, the function f is seen to be bijective so that the recovery of $\mathbb{E}[X] = \int_{\mathbb{R}} x\, dP_X(x)$ and the recovery P_X are one and the same.

Example 2.4 (Non-parametric Regression)**.** The problem of non-parametric regression as a descriptive problem is not always clearly defined in the literature. It can be properly formalized and interpreted by exhibiting its forward structure. Consider observations $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)) \in \mathbb{R}^{2n}$ assumed to be the realizations of some pairs of random variables $((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$ that are further assumed to be i.i.d. from the distribution $P_{X,Y}$ with marginals $P_X, P_Y \in L^2(\lambda)$ where λ is some dominating measure and L^2 is the Lebesgue space of square-integrable functions. The forward structure $(\mathscr{P}, f)$ for the problem can be directly stated for $P_{X,Y}$ in virtue of the i.i.d. assumption (see Remark 2.1). The possible descriptions of the phenomenon under study are thus restricted by selecting

$$\mathscr{P} = \left\{P_{A,B} \in \mathcal{M}(\mathbb{R}^2) : P_A, P_B \in L^2(\lambda)\right\},$$

where P_A and P_B denote the marginal for $P_{A,B}$. The restriction of the possible descriptions comes from a mild non-parametric assumption (square integrability) and an i.i.d. assumption. Once the set $\mathscr{P}$ of possible descriptions is chosen, the quantity and type of information to extract are restricted by considering the best L^2-approximation of Y as a function of X, that is,

$$\begin{aligned} f\colon \mathscr{P} &\to \left\{m \in \mathbb{R}^{\mathbb{R}} : m \text{ meas.}\right\}, \\ P_{A,B} &\mapsto \arg\min_{m \text{ meas.}} \mathbb{E}\left[(B - m(A))^2\right] \end{aligned}$$

where the random vector (A, B) has distribution $P_{A,B}$. In virtue of the space $\mathscr{P}$, the minimizer in the definition of f exists and is known to be the conditional expectation $m^*(A) := \mathbb{E}[B \mid A]$. However, the conditional expectation is not unique, so that f is not a function as defined above. To make it so, one solution is to change the definition of f by amending the codomain from measurable functions to equivalent classes of functions that only differ on non-negligible sets. If we do so, then f becomes a well-defined function, and the associated statistical forward problem becomes well-defined. The statistical problem then consists in estimating $f(P_{X,Y}) = \mathbb{E}[Y|X]$ from the observations $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$ that are assumed to be independent realizations of (X, Y). The choice of f is motivated by the fact that the joint distribution of two variables is hard to recover non-parametrically whereas the L^2-approximation (i.e., the conditional expectation) is a good enough measure of association that yields tractable inferential problems[2].

Example 2.5 (Linear Regression)**.** Linear regression is a problem that belongs to the same class of problems as non-parametric regression in the sense that they consider the same set of possible descriptions. However, the choice of f differs. The idea is to restrict attention to the best linear L^2-approximation of Y as a function of X among all such L^2-approximations. That is, the filter f is chosen as

$$\begin{aligned} f\colon \mathscr{P} &\to \left\{m \in \mathbb{R}^{\mathbb{R}} : m \text{ linear}\right\}, \\ P_{A,B} &\mapsto \arg \min_{m \text{ linear}} \mathbb{E}\left[(B - m(A))^2\right]. \end{aligned}$$

This is motivated by the fact that non-parametric regression may still be too hard an inferential problem so that additionally assuming linearity may present a better trade-off in terms of inferential difficulty and completeness of the description. By extending the definition of f to equivalence classes of functions as in non-parametric regression, f becomes a well-defined function and the forward problem is well-defined. However, it is common in the literature to slightly modify the problem for notational simplicity by considering the following filter

$$\bar{f}\colon P_{A,B} \mapsto \arg\min_{\beta \in \mathbb{R}} \mathbb{E}\left[(B - \beta A)^2\right].$$

While the solution is seen to equivalent to the previous problem, there is a subtlety in this reframing since the function $\bar{f}$ may not be well-defined anymore due to the fact that the system $\mathbb{E}[(B - \beta A)A] = 0$ may have several solutions. If we want to keep the codomain[3] of $\bar{f}$ as $\mathbb{R}$, then the set of possible descriptions can be changed to

$$\bar{\mathscr{P}} = \left\{P_{A,B} \in \mathcal{M}_2(\mathbb{R}^2) : P_A, P_B \in L^2(\lambda),\ \int_{\mathbb{R}} x^2 \, dP_A(x) \neq 0\right\}.$$

Indeed, the condition $\mathbb{E}[A^2] = \int x^2 \, dP_A(x) \neq 0$ then ensures that the minimizer in the definition of $\bar{f}$ is unique with closed-form expression $\beta^* = (\mathbb{E}[A^2])^{-1}\mathbb{E}[AB]$. The statistical problem then consists in estimating $\bar{f}(P_{X,Y}) = (\mathbb{E}[X^2])^{-1}\mathbb{E}[XY]$ from the observations $((x_1, y_1), (x_2, y_2), \dots, (x_n, y_n))$ that are assumed to be independent realizations of (X, Y).

[2] A weaker version of non-parametric regression can be considered by taking L^1 spaces in the definition of $\mathscr{P}$ and directly taking the conditional expectation operator in the definition of f. This generalization comes at the cost of the geometric interpretation of the problem.

[3] Another solution is to change the codomain of $\bar{f}$ from $\mathbb{R}$ to sets of $\mathbb{R}$, but it is rarely considered in practice.

2.2 Technical remarks

The nature and structure of forward problems is well-captured by the previous elementary examples. A number of more technical issues of varying importance are considered in this section. We make clear that:

- forward problems can be completely isolated from any stochastic model involving other distributions than the observable distributions – see Remark 2.2;
- forward problems induce statistical models in the sense of parametrized families of measures, but the forward structure logically precedes any parametrization – see Remark 2.3;
- there is no problem of identification in forward problems – see Remark 2.4;
- specification of forward problems in the sense that $P_{\mathbf{Z}} \in \mathscr{P}$ entirely captures the validity of the descriptions obtained from solving them – see Remark 2.5;
- the structure of forward problems offers a simple way to compare statistical models whose objects are descriptions – see Remark 2.6.

Remark 2.2 (Stochastic Model). A forward problem can be completely isolated from any stochastic model involving other distributions than the observable one because the observable distribution is directly taken as the object of inference. Inference then proceeds only on the assumption, consubstantial to any model-based inferential problem, that the observations are realizations of this distribution. The choice of a set $\mathscr{P}$ may be motivated by an underlying theory (see Remark 3.7) but the resulting forward problem can be solved as if this theory did not exist. Even more, theoretically at least, the set $\mathscr{P}$ could be taken so as to be without constraint on the observable distribution. It is this border case that makes the descriptive nature of forward problems explicit: because everything is assumed observed, the theoretical first-best cannot be anything more than the direct mapping procedure of finding the distribution that led to these observations.

Remark 2.3 (Statistical Model). A forward problem $(\mathscr{P}, f)$ for $P_{\mathbf{Z}}$ indirectly induces a statistical model for $P_{\mathbf{Z}}$ in the sense of a parametrized family of measures containing $P_{\mathbf{Z}}$. It is obtained by considering $\{P^\theta \in \mathscr{P} : \theta \in \Theta\}$ where $\Theta = f(\mathscr{P}) = E$ and $\theta_0 = f(P_{\mathbf{Z}})$. It is important to note, however, that the parametrization is not necessarily a valid parametrization for the original set $\mathscr{P}$ since the induced index relation $\theta = f(P) \mapsto P$ is not necessarily a valid mapping from Θ to $\mathscr{P}$ (as some image may have several preimages). It is so only when f is a bijective function (which is not guaranteed for many problems of interest). When f is not bijective, the parametrization implies the existence of an unindexed set $\mathscr{P}'$ of measures strictly contained in $\mathscr{P}$ for which the parametrization holds. It hence follows that $\{P^\theta \in \mathscr{P} : \theta \in \Theta\}$ may not be exactly the same structure as $(\mathscr{P}, f)$. However, the difference is mild since the statistical task of gaining knowledge on $f(P_{\mathbf{Z}}) = \theta_0$ remains identical. The main issue is different and follows from the fact that, in absence of a forward structure, it is unclear how to come up with well-behaved parametrized families like $\{P^\theta \in \mathscr{P}' : \theta \in \Theta\}$ for some potential set $\mathscr{P}'$ beyond simple problems with natural bijective parametrizations. We believe that, practically and logically, forward structures always precede indexed families in inferential problems for observable processes. It is then unclear if statistical models are needed when it is possible to exhibit a forward structure $(\mathscr{P}, f)$. (This point should be appreciated in light of the next remark concerning the problem of identification).

Remark 2.4 (Identification). The notion of identification can be defined for indexed statistical models where the object of inference is the indexing parameter: namely, an indexed statistical model $\{P^\theta : \theta \in \Theta\}$ with inferential object of interest θ is said to be identifiable if the map $\theta \mapsto P^\theta$ is injective. We show that an inferential problem with explicit forward structure $(\mathscr{P}, f)$ never leads to problems of identification. First, if $\mathscr{P}$ is not indexed, there is no problem of identification. If $\mathscr{P}$ is indexed, the index either plays no role or is selected as the object of inference through f, but if it is selected through the choice of f, it is de facto identified as the map $\theta \mapsto P^\theta$ is injective due to the fact that f has to be a well-defined function. This result can be alternatively obtained by considering the induced statistical model $\{P^\theta \in \mathscr{P} : \theta \in \Theta\}$ where $\Theta = f(\mathscr{P}) = E$ with implied unindexed set $\mathscr{P}' \subseteq \mathscr{P}$: the index map $I \colon \Theta \to \mathscr{P}', \theta = f(P) \mapsto P$ is always injective since f is a function and $\Theta = f(\mathscr{P}) = f(\mathscr{P}')$. The absence of identification issues in forward problems is the direct manifestation of the fact that forward problems, as their name suggests, do not call for the reconstruction of a partially observable object (as compared to problems introduced in Section 3).

Remark 2.5 (Specification). The problem of specification for forward problems relates to the choice of the initial set $\mathscr{P}$ of probability of measures, that is, the set of possible descriptions. A forward problem is said to be correctly specified if the condition $P_{\mathbf{Z}} \in \mathscr{P}$ is satisfied. The problem of specification in forward problems entirely captures the problem of the validity of the descriptions obtained from solving them: that is, if a forward problem $(\mathscr{P}, f)$ for $P_{\mathbf{Z}}$ is correctly specified, then uncertainty quantification around an estimate of $f(P_{\mathbf{Z}})$ can be interpreted as measuring the plausibility that the approximate description given by this estimate holds. The correctness of some specification is, however, rarely faced in a direct way for forward problems: the condition $P_{\mathbf{Z}} \in \mathscr{P}$ is an assumption that cannot be directly put to the test since it sustains the inferential task from which we try to gain knowledge of the unknown distribution $P_{\mathbf{Z}}$. One partial solution to gauge a specification is to split the observations into two sets where the specification's correctness is assessed on the first part and inference is performed on the second part. Beyond the problem of specification lies the more practical problem of deciding how much to restrict the set $\mathscr{P}$ so as to make the problem tractable without assigning a completely discordant distribution to the phenomenon. In general, there exists a clear trade-off between the cardinality of $\mathscr{P}$, the cardinality of $f(\mathscr{P})$, and the easiness of inference.

Remark 2.6 (Model Comparison). The forward structure has another advantage compared to parametrized statistical models: it allows for a simple way to compare problems in terms of inferential difficulty. This motivates two definitions that are illustrated in the previous examples.

- Two forward problems $(\mathscr{P}, f)$ and $(\mathscr{P}', f')$ are said to belong to the same class if $\mathscr{P} = \mathscr{P}'$. In other words, problems in the same class consider the same possible distributions for the observables but differ in terms of what parts of the distributions should be recovered as captured by different functions f and f'. All problems in the same class would be solved if we had knowledge of the unknown distribution of the observables. The difficulty of inference for problems in the same class then depends primarily on the complexity of $f(\mathscr{P})$ and $f'(\mathscr{P})$. This is illustrated by the problems of linear regression and non-parametric regressions of Examples 2.5 and 2.4 that are seen to belong to the same class (provided linear regression is stated with measurable functions as solutions).

- A forward problem $(\mathscr{P}', f')$ is said to be a sub-problem of another forward problem $(\mathscr{P}, f)$ if $\mathscr{P}' \subset \mathscr{P}$ and if f' is the restriction of f to $\mathscr{P}'$. In other words, a sub-problem considers a simpler inferential problem by restricting the set of possible distributions to be considered. The difficulty of inference is then directly captured by the complexity of the sets $\mathscr{P}$ and $\mathscr{P}'$. For instance, normal mean estimation as introduced in Example 2.3 is directly seen to be a sub-problem of non-parametric mean estimation as introduced in Example 2.2.

3 Explanation as inverse problems

3.1 Definition and examples

In relation to description, there exists a more elaborate task: explanation. The difference between description and explanation has led to a number of probabilistic and statistical considerations within philosophy (see, e.g., Woodward and Ross (2021)) but it has not been directly tackled within statistics. We introduced forward problems in Section 2 to capture the task of description in statistics. We argue in this section that there already exist problems in statistics that capture the task of explanation: inverse problems. The first step, however, is to be more precise since inverse problems have never been properly formalized in the literature or, if so, only as stochastic perturbations of deterministic inverse problems in the physical sciences. We thus introduce a simple structure that serves two purposes: it characterize all inverse problems encountered in practice and makes explicit their connection with the task of explanation.

The main idea is to view the recovery of an object that is only indirectly observable as the common denominator of inverse problems in statistics[4]. This exercise only makes sense because of the preliminary specification of a stochastic model tying observable processes to unobservable processes. Importantly, this model need not be exclusively derived from a deterministic inverse problem (see Remark 3.3 and Example 3.7). The model, however, commits the modeler to an explanation and confers to the indirectly observable object its meaning and a way for its recovery. Given a set of observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ that are assumed to be the realization of some random process $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ with joint distribution $P_{\mathbf{Z}}$, an inverse problem in statistics can thus be understood as the sequential choice of:

1. a stochastic model in the form of a functional equation $SM(P_{\mathbf{Z}}, P_{\mathbf{U}})$ where $P_{\mathbf{U}}$ is the joint

[4]This is already the intuition of Cavalier (2011) but this intuition is not directly formalized.

distribution of some unobservable random process $\mathbf{U}$: this corresponds to the choice of an explanation linking observable to unobservable processes.

2. a singleton-valued function $g\colon \{(P_{\mathbf{Z}}, P_{\mathbf{U}})\} \to \{g(P_{\mathbf{Z}}, P_{\mathbf{U}})\}$ where $P_{\mathbf{Z}}$ and $P_{\mathbf{U}}$ satisfy $\mathrm{SM}(P_{\mathbf{Z}}, P_{\mathbf{U}})$: this corresponds to the choice of the indirectly observable object whose meaning derives from the explanatory model.

The principle behind inverse problems is that by hypothesizing a functional equation between observable and unobservable processes, the modeler can recover some indirectly observable object. The model then commits the modeler to an explanatory model from which the object of interest gains its interpretation. The outcome is not simply a description, but an explanation of some observable phenomena by way of an underlying unobservable structure. The dependence of inverse problems to a theoretical bedrock is made clear in Subsection 3.3.

Definition 3.1 (Inverse Problem)**.** Consider a set of observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ assumed to be the realizations of a random process $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_n)$ with unknown joint distribution $P_{\mathbf{Z}}$. An inverse problem for $P_{\mathbf{Z}}$ is characterized by a triple $(\mathbf{U}, \mathrm{SM}(\cdot,\cdot), g)$ where $\mathbf{U}$ is an unobservable random process, $\mathrm{SM}(\cdot,\cdot)$ is a functional equation between the distribution $P_{\mathbf{U}}$ of $\mathbf{U}$ and the distribution $P_{\mathbf{Z}}$ of the observable process $\mathbf{Z}$, and g is a singleton-valued function. The statistical problem consists in gaining knowledge of $g(P_{\mathbf{Z}}, P_{\mathbf{U}})$ from the observations $\mathbf{z} = (z_1, z_2, \ldots, z_n)$.

As for forward problems, the formal definition of inverse problems hides the simplicity of the procedure under consideration. This simplicity will become clear as we consider a set of standard inverse problems. However, we need first to consider a fundamental difference with forward problems that follows from the presence of a stochastic model involving unobservable processes. This leads to a complication that is known as the problem of identification and that gives to inverse problems their name. The object $g(P_{\mathbf{Z}}, P_{\mathbf{U}})$ to be recovered depends on the distribution of some unobservable processes for which we have, by definition, no empirical knowledge. For the problem to be solvable, an additional step is needed – the identification step. That is, additional a priori assumptions on the distributions $P_{\mathbf{Z}}$ and $P_{\mathbf{U}}$ are needed so that a known function G can be found such that $g(P_{\mathbf{Z}}, P_{\mathbf{U}}) = G(P_{\mathbf{Z}})$. This step is what allows to reconstruct the object of interest from indirect observations of it. Since the reconstruction amounts to "undoing" the effect of the unobservable processes on the object of interest, it justifies calling these problems "inverse" (see Remark 3.3 for additional clarifications on the terminology).

Definition 3.2. An inverse problem $(\mathbf{U}, \mathrm{SM}(\cdot,\cdot), g)$ for $P_{\mathbf{Z}}$ is said to be identified if there exists a known function G such that $g(P_{\mathbf{Z}}, P_{\mathbf{U}}) = G(P_{\mathbf{Z}})$. The additional assumptions on $P_{\mathbf{Z}}$ and $P_{\mathbf{U}}$ needed to obtain G, if any, are called identification assumptions for G.

Identification as defined above does not correspond to the standard definition of identification in terms of the injectivty of an index map in a parametrized statistical model. The equivalence is proved in Remark 3.4 where we connect inverse problems with statistical models. To avoid too many abstract discussions at this stage, we instead demonstrate that inverse problems are very common in model-based inferential statistics. To illustrate this, we show that the forward problems considered in Examples 2.2, 2.3, 2.4, and 2.5 have mirror formulations as inverse problems. While these problems lead to exactly identical statistical procedures, their structure and interpretation differ significantly.

Remark 3.1. The functional equation and the object of interest in an inverse problem $(\mathbf{U}, \mathrm{SM}(\cdot,\cdot), g)$ are commonly expressed directly for the random processes – we should abuse language and use the notations SM and g to also denote functional equations and functions such as $\mathrm{SM}(\mathbf{Z}, \mathbf{U})$ and $g(\mathbf{Z}, \mathbf{U})$.

Example 3.3 (White Noise Model)**.** The forward problem of non-parametric mean estimation considered in Example 2.2 has a mirror formulation as an inverse problem with stochastic model known as the White Noise Model (due to its application in signal processing). Given some observations $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ of some observable random variables $(X_1, X_2, \dots, X_n)$, we consider the stochastic model given by

$$\mathrm{SM}(X_i, (\mu_0, \varepsilon_i)) \colon\; X_i = \mu_0 + \varepsilon_i,$$

for all $i \in \{1, 2, \dots, n\}$ where the processes (μ_0, ε_i)'s are unobservable with μ_0 constant in $\mathbb{R}$ and ε_i i.i.d. with common distribution $P_\varepsilon \in \mathcal{M}(\mathbb{R})$ such that $\mathbb{E}[\varepsilon] = \int_{\mathbb{R}} x \, dP_\varepsilon(x) = 0$. This works as an explanatory model: the observations are assumed to have been generated by the unknown true constant signal plus some i.i.d. unobservable white noise. The object of interest is then taken as

$$g(X_i, (\mu_0, \varepsilon_i)) = \mu_0$$

and is interpreted as the true constant signal. Since $\mu_0 = \mathbb{E}[X_1]$, the inverse problem is directly seen to be identified where G is simply the mean operator. The statistical problem is then the same as in Example 2.2 and consists in estimating $\mathbb{E}[X_1] = \mu_0$ from the observations $(x_1, x_2, \dots, x_n)$ that are assumed to be independent realizations of X_1 with distribution $P_{X_1} \in \mathcal{M}(\mathbb{R})$ with $\mathbb{E}[X_1] < \infty$.

Example 3.4 (Gaussian White Noise Model)**.** The forward problem of normal mean estimation considered in Example 2.3 has also a mirror formulation as an inverse problem with stochastic model known as the Gaussian White Noise Model. This problem is very similar to the white noise model: the stochastic model has the same functional formulation but the unobservable noises are now assumed to be i.i.d. Gaussian with mean zero and known variance. That is, given some observations $(x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ of observable random variables $(X_1, X_2, \dots, X_n)$, the stochastic model is now hypothesized as

$$\mathrm{SM}(X_i, (\mu_0, \varepsilon_i)) \colon\; X_i = \mu_0 + \varepsilon_i,$$

for all $i \in \{1, 2, \dots, n\}$ where the (μ_0, ε_i)'s are unobservable processes with μ_0 constant in $\mathbb{R}$ and ε_i i.i.d. with common distribution $N(0, 1)$. This works as an alternative explanatory model: the observations are now assumed to have been generated by the unknown true constant signal plus some i.i.d. unobservable Gaussian white noise. As in the white noise model, the object of interest is taken to be

$$g(X_i, (\mu_0, \varepsilon_i)) = \mu_0$$

and is interpreted as the true constant signal. Since $\mu_0 = \mathbb{E}[X_1]$, the inverse problem is directly seen to be identified where G is again the mean operator. The statistical problem is then the same as in Example 2.3 and consists in estimating $\mathbb{E}[X_1] = \mu_0$ from the observations $(x_1, x_2, \ldots, x_n)$ that can be seen as independent realizations of $X_1 \sim N(\mu_0, 1)$.

Example 3.5 (Non-parametric Regression Model)**.** The forward problem of non-parametric regression considered in Example 2.4 has a mirror formulation as an inverse problem. The two problems are often conflated in practice leading to many unfortunate confusions and interpretative mistakes. We clarify the formluation of non-parametric regression as an inverse problem. Given some observations $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)) \in \mathbb{R}^{2n}$ of some observable processes $((X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n))$, we consider the stochastic model given by

$$\mathrm{SM}((X_i, Y_i), (m_0, \varepsilon_i)) \colon\ Y_i = m_0(X_i) + \varepsilon_i,$$

for all $i \in \{1, 2, \ldots, n\}$ where the processes (m_0, ε_i)'s are unobservable, the variables (X_i, ε_i)'s are i.i.d. with common joint distribution in $\mathcal{M}(\mathbb{R}^2)$, and the process m_0 is constant with value in some functional space M. This works as an explanatory model: the observable variable Y_i is hypothesized to be the sum of some unobservable variable ε_i and a fixed unknown measurable function m_0 of the other observable variable X_i. The object of interest is then taken as

$$g((X_i, Y_i), (m_0, \varepsilon_i)) = m_0$$

and is interpreted as the functional relation between the Y_i's and the X_i's. The problem is, however, not identified as such. A standard set of identification assumptions[5] consists in hypothesizing that X_1, Y_1, and ε_1 lie in $L^2(\Omega, \mathcal{B}, \mathbb{P})$ for some probability space $(\Omega, \mathcal{B}, \mathbb{P})$ and that $\mathbb{E}[\varepsilon_1 | X_1] = 0$. This allows us to identify m_0 with G as the conditional expectation operator since it is seen from applying G to the stochastic model that $m_0 = \mathbb{E}[Y_1 | X_1]$. From a technical viewpoint, however, the object m_0 is not exactly identified because the operator G when defined on M has unique images only almost surely. To obtain full identification, the stochastic model has to be slightly amended by taking as image set the equivalent classes for null sets of functions in M. The statistical problem is then the same as in Example 2.4 and consists in estimating $\mathbb{E}[Y_1 | X_1] = m_0$ from the observations $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n))$ that can be seen as independent realizations of (X_1, Y_1).

Example 3.6 (Linear Regression Model)**.** Linear regression considered as a forward problem in Example 2.5 has also a dual formulation as an inverse problem that bears close resemblance to the non-parametric regression model. In this case, the stochastic model and the implied explanation

[5] A more traditional set of conditions consists in working with $L^1(\Omega, \mathcal{B}, \mathbb{P})$ instead of $L^2(\Omega, \mathcal{B}, \mathbb{P})$ as done in non-parametric regression as a forward problem. This is due to the fact that the geometric interpretation here is secondary given the stochastic model while the L^1 relaxation makes the identification assumption more plausible.

are changed to

$$\text{SM}((X_i, Y_i), (\beta_0), \varepsilon_i) \colon \ Y_i = \beta_0 X_i + \varepsilon_i$$

for all $i \in \{1, 2, \ldots, n\}$ where β_0 is constant with value in $\mathbb{R}$. For this problem (as in Example 2.5), the object to be recovered is directly taken to be β_0 (and not the measurable functions $x \mapsto \beta_0 x$ as in the non-parametric case), that is,

$$g((X_i, Y_i), (\beta_0, \varepsilon_i)) = \beta_0.$$

This problem is not identified as such. A standard set of identification assumptions for this problem is given by assuming that X_1, Y_1, ε_1 lie in $L^2(\Omega, \mathcal{B}, \mathbb{P})$ for some probability space $(\Omega, \mathcal{B}, \mathbb{P})$, that $\mathbb{E}[X_1 \varepsilon_1] = 0$, and that $\mathbb{E}[X_1^2] \neq 0$. It is then easily seen that the object of interest β_0 is identified by taking $G \colon P_{\mathbf{Z}} = (P_{(X_1,Y_1),\ldots,(X_n,Y_n)}) \to (\mathbb{E}[X_1^2])^{-1}\mathbb{E}[Y_1 X_1]$. The statistical problem is again the same as in Example 2.5 and consists in estimating $(\mathbb{E}[X_1^2])^{-1}\mathbb{E}[Y_1 X_1]$ from the observations $((x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n))$ that can be seen as independent realizations of (X_1, Y_1).

3.2 Technical remarks

The nature and structure of inverse problems in statistics are made explicit in the previous elementary examples. A number of more technical subtleties are covered in this section. In particular, we show that:

- inverse problems can target any features of the unobservable distributions – see Remark 3.2;
- inverse problems in statistics need not be connected to any inverse problem in operator theory – see Remark 3.3;
- inverse problems induce statistical models in the sense of parametrized families of measures that lead to the more common definition of identification, but the inverse structure logically precedes any such construction – see Remark 3.4;
- the separation of inverse problems into a preliminary stochastic model and an identification step is always arbitrary – see Remark 3.5;
- the problem of specification for inverse problems, which is more convoluted than for forward problems due to a stochastic model and identification assumptions, always makes the validity of the inferential results conditional on an unavoidable "leap of faith" – see Remark 3.6;
- identified inverse problems always induce forward problems – see Remark 3.7.

Remark 3.2 (Distributional Objects). In the examples above, the object of interest was chosen as a constant unobservable process: μ_0, m_0, and β_0. This need not always be the case. The object to be recovered in an inverse problem can be more generally chosen as any feature of the distribution of unobservable random processes that need not be constant. To illustrate this, consider in the problem of Example 3.4 a more general hypothesis for the explanation that consists in taking

$X_i = \nu_i + \varepsilon_i$ where the signals ν_i's are not assumed constant anymore but i.i.d. random variables independent of the ε_i's and with continuous support. In this case, the random process ν_i cannot be taken as an object to be recovered because it is not constant. However, it is natural to choose as inferential object its distribution or some of its moments. Another illustration of this more general procedure is considered in Example 3.7.

Remark 3.3 (Inverse Problems in Statistics and in Operator Theory). Inverse problems in statistics have been historically tied to deterministic inverse problems in the sciences involving the inversion of a mathematical operator (see Subsection 1.2). It is directly seen in the previous examples that, while all the problems could be interpreted as stochastic perturbations of deterministic systems, not all these deterministic systems generate inverse problems from the viewpoint of operator theory (only Example 3.6 and Example 3.5 do). Moreover, it will be seen in subsection 3.3 that not all measurement issues generating inverse problems in statistics originate from the stochastic perturbation of deterministic systems. This should make clear the difference of nature between inverse problems in statistics and inverse problems in operator theory in spite of the historical connection and the shared terminology.

Remark 3.4 (Statistical Model and Identification). As forward problems, inverse problems indirectly induce statistical models in the form of parametrized families of measures. However, these families are convoluted and play little role beyond the stochastic models used to define them. They have a tendency to hide the nature of inverse problems and are rarely, if ever, considered in practice. The rare cases when they are used are to introduce abstractly the problem of identification. For completeness, we show how to obtain them. From an inverse problem $(\mathbf{U}, \mathrm{SM}(\cdot,\cdot), g)$, we can define two statistical models: one for $P_{\mathbf{Z}}$ and one for the pair $(P_{\mathbf{Z}}, P_{\mathbf{U}})$. We start from unindexed family of measures for $P_{\mathbf{Z}}$ and $(P_{\mathbf{Z}}, P_{\mathbf{U}})$. If $\mathbf{Z}$ and $\mathbf{U}$ have values in some space E and F, respectively, then we can consider

$$\mathscr{P}_{\mathbf{Z}} = \{P_{\mathbf{A}} \in \mathcal{M}(E) : \exists P_{\mathbf{B}} \in \mathcal{M}(F) \text{ s.t. } \mathrm{SM}(P_{\mathbf{A}}, P_{\mathbf{B}}) \text{ holds}\}$$

and

$$\mathscr{P}_{\mathbf{Z},\mathbf{U}} = \{(P_{\mathbf{A}}, P_{\mathbf{B}}) \in \mathcal{M}(E) \times \mathcal{M}(F) : \mathrm{SM}(P_{\mathbf{A}}, P_{\mathbf{B}}) \text{ holds}\}.$$

The object of interest $\theta_0 = g(P_{\mathbf{Z}}, P_{\mathbf{U}})$ can then be used as parameter. This leads to two statistical models

$$\{P^{\theta} \in \mathscr{P}_{\mathbf{Z}} : \theta \in \Theta\}$$

and

$$\{P^{\theta} \in \mathscr{P}_{\mathbf{Z},\mathbf{U}} : \theta \in \Theta\}$$

where $\Theta = g(\mathscr{P}_{\mathbf{Z},\mathbf{U}})$ (where we abuse notation by extending the definition of g from singletons to sets). As for forward problems, the parametrization is not necessarily valid for the unindexed sets $\mathscr{P}_{\mathbf{Z}}$ and $\mathscr{P}_{\mathbf{Z},\mathbf{U}}$ since the function g may not be bijective. However, they remain valid for some sets $\mathscr{P}'_{\mathbf{Z}}$ and $\mathscr{P}'_{\mathbf{Z},\mathbf{U}}$ contained in $\mathscr{P}_{\mathbf{Z}}$ and $\mathscr{P}_{\mathbf{Z},\mathbf{U}}$, respectively. Identification then rewrites in terms of the injectivity of the index map for $\mathscr{P}'_{\mathbf{Z}}$ given by $I\colon \Theta = g(\mathscr{P}_{\mathbf{Z},\mathbf{U}}) \to \mathscr{P}'_{\mathbf{Z}}, \theta = g(P_{\mathbf{A}}, P_{\mathbf{B}}) \mapsto P_{\mathbf{A}}$. Indeed, it is not always true that I is injective: if $I(g(P_{\mathbf{A}}, P_{\mathbf{B}})) = I(g(P'_{\mathbf{A}}, P'_{\mathbf{B}}))$ so that $P_{\mathbf{A}} = P'_{\mathbf{A}}$, then it is not always true that $g(P_{\mathbf{A}}, P_{\mathbf{B}})) = g(P_{\mathbf{A}}, P'_{\mathbf{B}})$; the stochastic model may hold for different unobservable processes $\mathbf{B}$ and $\mathbf{B}'$. If the index map I is injective, however, then it is invertible (due to the restriction to $\mathscr{P}'$) and we have

$$I^{-1}(P_{\mathbf{A}}) = I^{-1}(I(g(P_{\mathbf{A}}, P_{\mathbf{B}}))) = g(P_{\mathbf{A}}, P_{\mathbf{B}}).$$

This is equivalent to the definition of identification given in Definition 3.2 where the function G is simply taken to be I^{-1}. This condition in terms of injectivity is often taken as the definition of identification. It is, however, never linked to practical inverse problems due to the difficulty to properly define statistical models for these problems. The difficulty is handled in the construction above but it also shows, by contrast, the logical primitivity and simplicity of the inverse structure exhibited in Definition 3.1 and of the characterization of identification given in Definition 3.2.

Remark 3.5 (Stochastic Model and Identification). An inconvenient fact of inverse problems is that there is no explicit difference between assumptions used for the definition of the stochastic model and assumptions used for the identification step. This is illustrated in Examples 3.3, 3.4, 3.5, and 3.6. In Examples 3.3 and 3.4, sufficiently many assumptions on $P_{\mathbf{Z}}$ and $P_{\mathbf{U}}$ are used in the definition of the stochastic model to make the identification step disappear. In Examples 3.5 and 3.6, the stochastic model has not been sufficiently characterized leading to a non-empty set of identification assumptions. For any given inverse problem, it is always possible to redefine the stochastic model by including the identification assumptions in it so as to make the identification step disappear. Conversely, it is always possible to strip a stochastic model of an already identified inverse problem from its assumptions to make the object of interest no longer identified. This reverse procedure can always be used as an informal way to determine where the identification step should start. It is important to remember, however, that there is always an arbitrariness in inverse problems as to when the stochastic model ends and the identification problem starts.

Remark 3.6 (Specification). The problem of specification in inverse problems relates to the choice of a stochastic model and of identification assumptions for the indirectly observable objects of interest. This implies, in particular, the hypothesis that there exist some unobservable processes $\mathbf{U}$ with distribution $P_{\mathbf{U}}$ that explain the observable processes $\mathbf{Z}$. An inverse problem is then said to be correctly specified if all the assumptions on $P_{\mathbf{Z}}$ and $P_{\mathbf{U}}$ implied by the stochastic model and the identification step hold for the phenomena under consideration. Given the structure of inverse problems, they naturally involve more assumptions than forward problems do, hence making specification in inverse problems a more challenging issue. However, there is a difference of even more importance. Because the assumptions in inverse problems bear on unobservable processes, they always stand beyond any possible empirical verification. This fundamental limit is qualitatively different than the limit we exhibited in forward problem (see Remark 2.5). If, in forward problems, the problem of specification is problematic as it sustains the inferential task, it is not impossible to gauge it by considering, say, splitting the observations into two sets: one for model confirmation, one for inference. This is not possible in inverse problems since there is, by definition, no observations for the unobservable processes. This creates an unavoidable "leap of faith" in inverse problems that stand beyond any empirical test. The validity of the inferential results for the phenomena under study are always conditional on this "leap of faith". This is the direct cost to bear for the theoretical interpretation of the inferred object.

Remark 3.7 (Inverse Problem as Forward Problem). An identified inverse problem always induce a forward problem with structure $(\mathscr{P}, f)$ where $\mathscr{P}$ is the unindexed family of measures for $P_{\mathbf{Z}}$ described in Remark 3.4 (where the identification assumptions are added to the stochastic model) and f is given by the closed-form identification function G. This connection, which holds conditional on identification, is another way to understand the qualitative difference between forward problems and inverse problems.

3.3 Theory-ladenness of inverse problems

The structure of inverse problems is noticeably different from the structure of forward problems. The specification of a stochastic model in terms of unobservable processes commit the modeler to more than a simple set of distributions for the observable distributions. In particular, prior observations of the observable phenomena cannot be the sole justification behind some inverse structure (see Remark 2.5). This difference begs the question as to where inverse problems originate.

The simple answer to that question is that a domain-specific theory is always preliminary to the formulation of an inverse problem. It is such a theory that guides and constrains the selection of a stochastic model for some unobservable process and the choice of the indirectly observable object to be recovered. It is the hypothesized probabilistic implications of the theory that are used to assess the plausibility of the assumptions used for identification of the indirectly observable object. This is what we refer to as the theory-ladenness of inverse problems. This theory-dependence is also the reason why inverse problems can be associated with the task of explanation: the stochastic model $\mathrm{SM}(\cdot, \cdot)$ and the unobservable process $\mathbf{U}$ gain their meaning from the meaning associated with some theory's model and unobservable entities. This can be illustrated with the white noise model of Example 3.3. As its name suggests, the white noise model is often used in a signal processing context to deconvolve a signal from noisy observations ot it. It is the physical theory behind signals and their measurements that then justify the inverse problem of Example 3.3 and give to the unobservable processes their meaning as signal and noise. The mean zero assumption for the noise is then typically motivated in virtue of the inherent symmetry of physical noise. The additional assumption of Gaussianity of the noise in the Gaussian white noise model of Example 3.4 is typically justified in this context by the physicist's knowledge of the random behavior of charged particles and their impact to the measurement device. The white noise model and its Gaussian restriction are, however, not reserved to signal processing: a number of theories from other fields can be invoked to justify the recovery of signal-like objects from noise-like processes.

The white noise model in the context of signal processing is a typical example of a deterministic physical inverse problem that is assumed to be perturbed by some unobservable stochastic noise. This type of problems has been traditionally the focus point of statisticians working with inverse problems (see Subsection 1.2). It is important to note, however, that such problems are not the only ones that fit into the category of statistical inverse problems defined in Section 3. Many problems, for which the measurement issue does not come from some unobservable stochastic disturbance of a deterministic model, can be considered. Examples of such inverse problems abound in the social sciences where selection issues emerge. Another typical example is given by problems of causal inference where the measurement issue comes from the fundamental problem of causal inference, that is, the fact that it is not possible to observe simultaneously the outcome with and without treatment for a given unit. What is important to note is that all these problems, even if they are not obtained from some deterministic inverse problems, are derived from a domain-specific theory and can be construed as explanatory models for the observable processes. This is illustrated in a concluding example where we look at model-based causal inference with potential outcomes. In this case, the domain-specific theory comes from a formalization of causality in terms of the theoretical notion of potential outcomes.

Example 3.7 (Model-Based Causal Inference with Potential Outcomes)**.** There exist different formalizations of causal inference. We consider here one formalization based on the notion of potential outcomes and resort to model-based inference[6]. For the sake of simplicity, we consider a treatment that can only be received or not. We denote the treatment indicator by $d \in \{0, 1\}$ where $d = 1$ means the treatment is received. The objective is to investigate the causal effect of the treatment on some real-valued outcome $y \in \mathbb{R}$. To capture causality, we introduce potential outcomes $y(1)$ and $y(0)$ that correspond to what the outcome would be if the treatment were to be received or not, respectively. The logic behind potential outcomes is that, when $d = 1$, the quantity $y(1)$ is equal to y while $y(0)$ cannot be known, and inversely when $d = 0$. This is the theoretical backdrop on which a statistical problem can be built. Given observed outcomes and observed treatment assignments $(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ and $(d_1, d_2, \ldots, d_n) \in \{0, 1\}^n$ for n different units, we theoretically have $2n$ potential outcomes $(y_1(1), y_2(1), \ldots, y_n(1)) \in \mathbb{R}^n$ and $(y_1(0), y_2(0), \ldots, y_n(0)) \in \mathbb{R}^n$ among which only n can be simultaneously known as given by the values $d_1, d_2, \ldots, d_n$. This is the basis of a model-based inferential problem by construing these outcomes, potential outcomes, and treatment assignments as the realizations of some random processes $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$, $\mathbf{Y}(1) = (Y_1(1), Y_2(1), \ldots, Y_n(1))$, $\mathbf{Y}(0) = (Y_1(0), Y_2(0), \ldots, Y_n(0))$, and $\mathbf{D} = (D_1, D_2, \ldots, D_n)$. We make a standard hypothesis by assuming the quadruples $(Y_i, Y_i(1), Y_i(0), D_i)$, $i = 1, 2, \ldots, n$, to be i.i.d. with common distribution that of a representative quadruple $(Y, Y(1), Y(0), D)$. A typical object of interest is then defined in terms of the common distribution of the potential outcomes variables. We consider here the (population) average treatment effect ((P)ATE) defined as $\mathbb{E}\left[Y(1) - Y(0)\right]$ and whose causal interpretation comes from the theoretical meaning a priori

[6]This is also known as the potential-outcome super-population approach to causal inference. Other approaches exist and not all of them lead to inverse problems nor to model-based inferential problems. In particular, with knowledge of the treatment assignment mechanism, causal inference is possible by only exploiting this knowledge – this is referred to as the finite-sample approach to causal inference and falls into the category of design-based inferential problems.

assigned to potential outcomes. We now show that this defines a standard inverse problem by exhibiting its inverse structure. The stochastic model is given by

$$\mathrm{SM}(Y_i, Y_i(1), Y_i(0), D_i) : Y_i = D_i Y_i(1) + (1 - D_i) Y_i(0)$$

for all $i \in \{1, 2, \ldots, n\}$. This works as an explanation for the observable outcomes $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$. It is directly motivated by the theory characterizing potential outcomes and the associated deterministic rule $y = dy(1) + (1 - d)y(0)$. The observational issue, however, does not emerge from the injection of some unobservable stochastic noise, but from the nature itself of potential outcomes. Indeed, the potential outcomes are theoretically unobservable entities which are naturally construed as unobservable processes in the inverse problem, that is,

$$\mathbf{U} = (\mathbf{Y}(1), \mathbf{Y}(0)) = ((Y_1(1), Y_2(1), \ldots, Y_n(1)), (Y_1(0), Y_2(0), \ldots, Y_n(0))).$$

The object of interest is the (P)ATE which rewrites as

$$g(P_{\mathbf{Y},\mathbf{D}}, P_{\mathbf{U}}) = \mathbb{E}\left[Y(1) - Y(0)\right].$$

The (P)ATE is not directly observable since $\mathbf{Y}(1)$ and $\mathbf{Y}(0)$ are unobservable processes. The recovery of this object is made possible from the stochastic model and additional identification assumptions. The most elementary identification assumption is given by

$$Y(0), Y(1) \perp\!\!\!\perp D$$

and allows to take

$$G(P_{\mathbf{Y},\mathbf{D}}) = \mathbb{E}\left[Y|D = 1\right] - \mathbb{E}\left[Y|D = 0\right]$$

as identifying function. This identification assumption naturally takes a particular interpretation in light of the nature of potential outcomes. It means that the treatment is randomly assigned independently of what the outcomes with and without treatment could be. This assumption cannot be verified empirically in virtue of the unobservable quality of potential outcomes – it has to be believed in a "leap of faith" (see Remark 3.6). It is typically believed to hold when the modeler has direct control over the treatment assignment mechanism and the units have no agency with respect to the treatment. This identification assumption is not the only one that can be considered for the problem. It can be easily relaxed to conditional independence and a support assumption if an extra variable is available. We refer, however, the reader to Ding (2024) for additional results on this problem. The simple case introduced above is enough to demonstrate the inverse nature of model-based causal inference problems and illustrate the theory-ladenness of statistical inverse problems beyond deterministic inverse problems in the sciences.

4 Conclusion

If one adventurously decides to visit the foundations of statistics, one is likely to encounter the remnants of philosophical musings on the nature of inference and the interpretation of probabilities. It is hard to find among these musings a set of explicit interpretative guidelines telling with clarity how to come up with a statistical model and how to interpret it beyond the general goal of "induction" and "information extraction". It has been the objective of this paper to make clear the assumptions that sustain model-based inferential problems and the type of knowledge that one gains when solving these problems. For this purpose, we introduced two simple structures that miror each other and generate most parametrized statistical models in practice. One structure can be associated with the task of description where a probability measure is simply assigned to some realizations of it. The other structure can be associated with the task of explanation as it allows for the recovery of objects only partially observable in virtue of a preliminary theoretical model. This second structure is seen to exactly generate what are traditionally referred as inverse problems in statistics and thus solves *en passant* a long standing issue in the foundations of statistics.

References

Aster, R. C., B. Borchers, and C. H. Thurber (2018): *Parameter Estimation and Inverse Problems*. Elsevier.

Bahadur, R. R., and L. J. Savage (1956): "The Nonexistence of Certain Statistical Procedures in Nonparametric Problems," *The Annals of Mathematical Statistics*, 27(4), 1115–1122.

Berger, J. O. (1985): *Statistical Decision Theory and Bayesian Analysis*. Springer New York.

Breiman, L. (2001): "Statistical modeling: The two cultures (with comments and a rejoinder by the author)," *Statistical Science*, 16(3), 199–231.

Cavalier, L. (2011): "Inverse Problems in Statistics," in *Inverse Problems and High-Dimensional Estimation: Stats in the Château Summer School, August 31-September 4, 2009*, pp. 3–96. Springer.

Ding, P. (2024): *A First Course in Causal Inference*. CRC Press.

Efron, B., and T. Hastie (2021): *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge University Press.

Evans, S. N., and P. B. Stark (2002): "Inverse problems as statistics," *Inverse Problems*, 18(4), R55.

Ferguson, T. S. (1967): *Mathematical Statistics: A Decision Theoretic Approach*. Academic Press.

Florens, J.-P., and A. Simoni (2017): "Introduction to the Special Issue on Inverse Problems in Econometrics," *Annals of Economics and Statistics/Annales d'Économie et de Statistique*, 128, 1–3.

Giné, E., and R. Nickl (2021): *Mathematical Foundations of Infinite-dimensional Statistical Models*. Cambridge University Press.

Lehmann, E. L., and G. Casella (1998): *Theory of Point Estimation*. Springer.

Lehmann, E. L., and J. P. Romano (2005): *Testing Statistical Hypotheses*. Springer.

Robert, C. P. (2007): *The Bayesian Choice: From Decision-theoretic Foundations to Computational Implementation*. Springer.

Shao, J. (2008): *Mathematical Statistics*. Springer.

Sudakovand, V., and L. Khalfin (1964): "Statistical approach to ill-posed problems in mathematical geophysics," *Sov. Math.—Dokl*, 157.

Tarantola, A. (2005): *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM.

van der Vaart, A. W. (2000): *Asymptotic Statistics*. Cambridge University Press.

Woodward, J., and L. Ross (2021): "Scientific Explanation," in *The Stanford Encyclopedia of Philosophy*, ed. by E. N. Zalta. Metaphysics Research Lab, Stanford University, Summer 2021 edn.