２０２５年度

大学院文学研究科博士課程前期２年の課程入学試験

（　夏期・一般選抜　）問題

専門科目　行動科学　専攻分野

試験開始の合図があるまで、この問題冊子を開いてはいけない。

受験記号番号	

成績	

2025年度

大学院文学研究科博士課程前期２年の課程入学試験

（夏期・一般選抜）問題

専門科目（　　行動科学　　専攻分野）

注意）解答の順序は自由であるが，どの問題の解答であるかが分かるように，問題番号を間違いなく記入すること．

問題 1-1.

ある大学入学試験において，受験者は男性が 60%を占めていた．合格率は，男性は男性受験者の 40%，女性は女性受験者の 30%であった．

(1)　この大学の男女合わせた全体の受験者に対する男性の合格率，女性の合格率を求めよ．
(2)　この大学の男女合わせた全体の合格率を求めよ．
(3)　この大学の合格者の名簿から一人をランダムに選んだ場合，それが女性である確率を求めよ．

問題 1-2.

2 つの確率変数 X, Yに対して，Xと$Y - E[Y|X]$が無相関になることを示せ．

（次頁に続く）

問題 2-1. 「1 時点の観察データから因果関係を検証することは難しいが，パネルデータを利用すれば可能である」という主張について，この主張が正しいと思うか，誤っていると思うかをまず選択し，その理由を述べなさい．肯定する場合でも否定する場合でも，追加的な条件が必要であるならば，その条件も明示しなさい．

問題 2-2. 確率変数 X と確率変数の組（Y, Z）が独立である．すなわち同時確率密度関数について

$$f(x,y,z) = f(x)f(y,z)$$

であるとき，「X と Y が独立である」かつ「X と Z が独立である」が成立することを証明しなさい．

（次頁に続く）

受験記号番号	

問題3. 次の語句について，1語句につき100字程度で簡潔に説明せよ.

①中範囲の理論　②疑似決定係数

問題4. 以下の英文を読み，問いに答えなさい.

(1) 下線部 (a) を日本語に訳しなさい.

(2) 下線部 (b) を日本語に訳しなさい.

(3) causal inference において著者が述べる"the second task"と"the third task"の違いについて，日本語で説明しなさい.

Causality is a very intuitive notion that is difficult to make precise without lapsing into tautology. Two ingredients are central to any definition: (1) a set of possible outcomes (counterfactuals) generated by a function of a set of "factors" or "determinants" and (2) a manipulation where one (or more) of the "factors" or "determinants" is changed. An effect is realized as a change in the argument of a stable function that produces the same change in the outcome for a class of interventions that change the "factors" by the same amount. The outcomes are compared at different levels of the factors or generating variables. <u>(a) Holding all factors save one at a constant level, the change in the outcome associated with manipulation of the varied factor is called a causal effect of the manipulated factor.</u> This definition, or some version of it, goes back to Mill (1848) and Marshall (1890). Haavelmo's (1943) made it more precise within the context of linear equations models. The phrase 'ceteris paribus' (everything else held constant) is a mainstay of economic analysis and captures the essential idea underlying causal models. This paper develops the scientific model of causality developed in economics and compares it to methods advocated in epidemiology, statistics, and in many of the social sciences outside of economics that have been influenced by statistics and epidemiology.

I make two main points that are firmly anchored in the econometric tradition. The first is that causality is a property of a model of hypotheticals. A fully articulated model of the phenomena being studied precisely defines hypothetical or counterfactual states. A definition of causality drops out of a fully articulated model as an automatic by-product. A model is a set of possible counterfactual worlds constructed under some rules. The rules may be the laws of physics, the consequences of utility maximization, or the rules governing social interactions, to take only three of many possible examples. A model is in the mind. As a consequence, causality is in the mind.

In order to be precise, counterfactual statements must be made within a precisely stated model. <u>(b) Ambiguity in model specification implies ambiguity in the definition of counterfactuals and hence of the notion of causality.</u> The more complete the model of counterfactuals, the more precise the definition of causality. The ambiguity and controversy surrounding discussions of causal models are consequences of analysts wanting something for nothing: a definition of causality without a clearly articulated model of the phenomenon being described (i.e., a model of counterfactuals). They want to describe a phenomenon as being modeled "causally" without producing a clear model of how the phenomenon being described is generated or what mechanisms select the counterfactuals that are observed in hypothetical or real samples. In the words of Holland (1986), they want to model the effects of causes without modeling the causes of effects. Science is all about constructing models of the causes of effects. This paper develops the scientific model of causality and shows its value in analyzing policy problems.

My second main point is that the existing literature on "causal inference" in statistics confuses three distinct tasks that need to be carefully distinguished:

- Definitions of counterfactuals.
- Identification of causal models from population distributions (infinite samples without any sampling variation). The hypothetical populations producing these distributions may be subject to selection bias, attrition, and the like. However, issues of sampling variability of empirical distributions are irrelevant for the analysis of this problem.
- Identification of causal models from actual data, where sampling variability is an issue. This analysis recognizes the difference between empirical distributions based on sampled data and population distributions generating the data.

(次頁に続く)

Table 1 represents these three tasks.

The first task is a matter of science, logic, and imagination. It is also partly a matter of convention. A model of counterfactuals is more widely accepted the more widely accepted are its ingredients, which are

- the rules of the derivation of a model including whether or not the rules of logic and mathematics are followed;
- its agreement with other theories; and
- its agreement with the accepted interpretations of facts.

Models are not empirical statements or descriptions of actual worlds. They are descriptions of hypothetical worlds obtained by varying—hypothetically—the factors determining outcomes.

TABLE 1
Three Distinct Tasks Arising from Analysis of Causal Models

Task	Description	Requirements
1	Defining the Set of Hypotheticals or Counterfactuals	A Scientific Theory
2	Identifying Parameters (Causal or Otherwise) from Hypothetical Population Data	Mathematical Analysis of Point or Set Identification
3	Identifying Parameters from Real Data	Estimation and Testing Theory

The second task is one of inference in very large samples. Can we recover counterfactuals (or means or distributions of counterfactuals) from data that are free of sampling variation? This is the identification problem. It abstracts from any variability in estimates due to sampling variation. It is strictly an issue of finding unique mappings from population distributions, population moments or other population measures to causal parameters.

The third task is one of inference in practice. Can one recover a given model or the desired causal parameters from a given set of data? This entails issues of inference and testing in real world samples. This is the task most familiar to statisticians and empirical social scientists. This essay focuses on the first two tasks. Identification is discussed, but issues of sampling distributions of estimators, such as efficiency, are not.

（出典） Heckman, James J. 2005. "The Scientific Model of Causality." *Sociological Methodology* 35(1): 1-97.

受験記号番号

受験記号番号	

受験記号番号	