Journal of Statistical and Mathematical Sciences
ISSN 3065-100X
Volume 12 Issue 3
July-September 2024
Impact Factor: 7.82
Enquiry: contact@continentalpub.online
Published by Continental Publication | http://continentalpub.online/index.php/jsms

MATHEMATICAL MODELS IN SCIENCE: BRIDGING DIFFERENCES AND DISCOVERING COMMONALITIES

Erik Magnus Johansen
Department of Statistics, University of Oslo, P.O.Box 1053 Blindern, N-0316 Oslo, Norway

Abstract: *Various scientific disciplines, including biometrics, chemometrics, psychometrics, and econometrics, share a deep connection with statistics but often employ specialized applications. Effective communication across these interdisciplinary boundaries necessitates a common language, and mathematics serves as a universal bridge among the natural sciences. This article centers on the language of mathematics and its role in scientific modeling. The thesis posited here is that mathematical models should be formulated in a manner comprehensible, with some effort, to serious researchers across all disciplines, despite potential cultural barriers. The notion of "understanding" is multifaceted, exemplified by the enigmatic realm of quantum mechanics. While quantum theory is traditionally presented in an abstract language that facilitates calculations, its interpretation remains contentious, posing challenges for researchers outside the physics domain. Notably, quantum mechanics was initially expressed in distinct languages—Schrödinger's wave mechanics and Heisenberg's matrix mechanics—before being mathematically unified. This unification yielded a mathematically elegant yet conceptually intricate theory, accompanied by over 16 partially conflicting interpretations. In the works of Helland (2021) and Helland (2022a), a novel foundation is proposed, potentially more accessible to researchers beyond the quantum community. It introduces conceptual variables denoted as θ, emerging from inquiries like "What will θ be if measured?" Particularly in the discrete case, often central in textbooks and research, precise responses in the form of "θ = u" can be obtained.*

Keywords: *Mathematics, interdisciplinary communication, mathematical models, quantum mechanics, conceptual variables.*

1 Introduction

Biometrics, chemometrics, psychometrics and econometrics are sciences that are closely connected to statistics but are more focused on specific applications. Communication across scientific borders is important, and to some extent such a communication requires a common language. The language of mathematics is common to all natural sciences. The present article will focus on this language, and on the use of mathematical models in science.

Even though there are cultural barriers, my thesis is that one should be able to formulate mathematical models in such a way that they - at least after some effort – could be understood by all serious researchers, in all disciplines. The word ‘understand’ is quite demanding. My late colleague Emil Spjøtvold once remarked that things can be understood on many different levels. This applies also in this case. Take one important example, quantum mechanics. In textbooks and in research papers, quantum theory is formulated in a very abstract language, and using this language, physicists can compute, but the issue of interpretation is very controversial, and researchers from other disciplines have great difficulties of understanding what is going on.

DOI: https://doi.org/10.5281/zenodo.14513487

In fact, quantum mechanics was originally formulated in two different languages, the wave mechanics of Schrödinger and the matrix mechanics of Heisenberg. By mathematical abstraction these two were unified, and that resulted in a mathematically clean, but conceptual difficult theory, which has been extremely difficult to interpret. The relevant Wikipedia article contains more than 16, partially mutually excluding, interpretations. In the book Helland (2021) and the article Helland (2022a), a new basis is proposed, a basis which is hopefully easier to understand for researchers outside the quantum community. As I see it, we can take as our point of departure a set of conceptual variables □in the mind(s) of the investigator(s), and these variables are introduced through questions of the form 'What will □be if we measure it?'. In the discrete case, which is the focus of many textbooks and research papers, one can obtain sharp answers of the form '□= u'.
For certain well-defined conceptual variables, these questions and answers together define the relevant quantum states. This should be compared to the ordinary quantum language, where a state is defined as a unit vector with arbitrary phase in an abstract, complex Hilbert space.
From this perspective I will discuss the concept of a mathematical model in general. I will cover several disciplines, but the story starts with the interplay between chemometrics and statistics. Since so many different areas are covered, I will attempt to be brief on each single discipline. This may imply some simplifications. In particular, the choice of references here is somewhat subjective. The important set of deterministic models will not be covered here. Nor will I mention stochastic models in, say, population dynamics and genetics. I will concentrate on models that, in some way of other, may be associated with areas connected to statistical inference.
The plan of the paper is as follows: In Section 2 I give a general discussion of cultures within science. Then, as a starting point, mathematical models used in statistical inference are described in Section 3. The Partial Least Squares 'model' as developed in Chemo metrics in the 1980's is discussed in Section 4, while Section 5 gives a link between Section 3 and Section 4. Dennis Cook's envelope model, which can be closely associated with this, is considered in Section 6. In Section 7 a brief discussion of the connection between statistical models and machine learning is given, while Section 8contrasts briefly statistical models and models connected to causality. Decisions and measurements are important both in connection to statistics and in connection to the foundation of quantum theory, and models for these are introduced in Section 9. Then in Section 10 the elements of a foundation of quantum mechanics, taking the model notionas a basis, are described. Section 11 gives some general discussion points.

2 On science and cultures within science

My definite starting point is that science is important in society. There are extremely many problems facing the world now: climate, health problems, poverty problems, refugee problems, international conflicts, and the existence of very dangerous weapons. For all these problems one should try to device rational solutions that at the same time satisfy good ethical standards. Rational and good ethical decisions should be made by national and international leaders. Unfortunately, in the process of making decisions, we may all be limited. (In a concrete physical situation where communication is important, this has been argued for in Helland, 2022b.) So ideally, to arrive at good decisions, in addition to other factors, our political leaders should attempt to get good, rational inputs. Ideally again, such inputs could in concrete cases be given by a joint effort by groups of scientists.
In many cases this requires time, and it may also meet practical difficulties. But it is important that scientists are available in such cases and have tools for obtaining the necessary scientific knowledge. In choosing tools, a scientist must make choices that sometimes are uncertain. It is then important that he can get inspiration from colleagues. Such set of colleagues may then be a part of a larger community. What we do not often think about, is that such communities form a joint culture, and that elements of such a culture may have arbitrary or historical roots. Sometimes a scientist's choice of tools may depend on what inputs he has obtained from his culture.

DOI: https://doi.org/10.5281/zenodo.14513487

An example can be the choice of regression method in the case of collinear data. Statisticians has many tools here, variable selection, ridge regression, principal component regression and so on. In addition, chemometricians have developed their own tool, partial least squares (PLS), which empirically seems to have good properties, and has now been applied in a great variety of fields (Mehmood and Ahmed, 2015).

What is culture? According to the author and philosopher Ralph D. Stacey it is a set of attitudes, opinions, and convictions that a group of people share, about how one should act towards each other, how things should be evaluated and done, which questions that are important and which answers that may be accepted. The most important elements in a culture are unconscious and cannot be forced upon from the outside.

3 Models, probabilities, and procedures in statistics

Most of what we do in statistics is based on probability models for the data, but there are exceptions (Breiman, 2001). I will disregard nonparametric models here, and assume continuous data, so that the model in general is given by a probability density f□(x); (1) where x is the data, and □is the full parameter. It is important that the model is selected by the researcher. The parameter □exists in some fundamental way in the researcher's mind, and is related to the research question that he wants to investigate. More precisely, the research question is often posed in terms of a sub-parameter □=□(□).

It is well known that statistical methods may be classified into frequents methods, Bayesian methods, and - as a perhaps unusual way of thinking - fiducial methods. To some extent this divides statisticians into different schools, but firstly, there are connections between the different set of methods, and secondly, one single person may also use methods from different schools.

The interpretation and use of the probability concept varies somewhat between different schools. The Bayesians have their prior distributions, and, from this and the statistical model, posterior distributions. A frequentist will sometimes distinguish between aleatoric probabilities - probabilities of data, given some assumed properties of the world - and epistemic probabilities - probabilities of some properties of the world, perhaps given some data. I will come back to the latter below. The former are the bases for constructing statistical models.

From a Bayesian perspective, everything can be modeled by using ordinary probability models. The priors and the posteriors both follow Kolmogorov's axioms and allother probability rules that we are used to. In particular, the posterior distributions are epistemic distributions that can be seen as ordinary probability distributions. In Schweder and Hjort (2016), confidence distributions were proposed from a frequentist point of view as a method of making 'probability' statements about aspects of the world, given some data from an experiment. Here is the definition from Schweder and Hjort (2016):

Definition 1 *A nondecreasing right-continuous function of the one-dimensional* □*,depending on the data Y, say C(*□*;Y), is the cumulative distribution function for a confidence distribution for* □*provided that it has a uniform distribution under the statistical model, whatever the true value of the full parameter* □.

More concretely, this can be obtained as follows: Take a class of one-sided confidence intervals from minus infinity to□□, corresponding to the confidence coefficient □. Then, we can find C from □□= □□(Y) =C^{-1}(□), so that C also depends on the data Y:

P□[□ ≤□□□] = P□[C(□) ≤C(□□)] = P□[C(□)≤□] = □ (2)

For a scalar continuous parameter, Fisher's fiducial distribution is essentially thesame as the confidence distribution. More generally, a fiducial distribution can be found (Taraldsen and Lindqvist, 2021) from a data generating model y=y(□;u), where□= □(□) is a parameter, and u has a probability

DOI: https://doi.org/10.5281/zenodo.14513487

distribution depending upon □. The model is called simple if it has a unique solution □= □(y;u). The fiducial distribution is then the distribution of □for fixed y.

Fisher's difficulties began when he tried to generalize his concept to the multivariate case. In the last few years, the concept of a fiducial theory has been taken up again, see references in Taraldsen and Lindqvist (2021). Older results by Fraser (1964) have tied the fiducial distribution to a Bayesian distribution with a group-invariant prior. A concept due to Fisher (1973) is the likelihood function L the probability density (1), seen as a function of q. Maximum likelihood estimation is used in many applied fields. Less well known is the likelihood principle, saying that observations with the same, or proportional, likelihood always give the same experimental evidence. In Helland (2021) I advocate a version where the experimental context is fixed, which in my opinion makes the principle less controversial. A well-known statistical model with many applications is the regression model

y = □+X□+e; (3)

where y is an n-dimensional vector of response variables, X is annxp matrix of explanatory variables, and e is a

random error term. When there is no collinearity, the maximum likelihood estimator under normality is

^{1}X'y: (4)

(X'X)

When there is collinearity, in particular, ifp >n, many different estimators have been proposed by statisticians. In general, our purpose might be to use this estimate in prediction of a newy-value y_0 from a set of x-values x_0.

4 A certain 'model' from chemometrics

During the 1980's chemometricians developed an algorithm which was intended to solve the above regression problem in cases with collinearity: The partial least squares, or PLS algorithm. The algorithm used had its origin in Herman Wold's general system analysis (Jöreskog and Wold, 1982). Briefly, a theoretical version of it goes as follows, where the starting point is an x-vector x (the column vector corresponding to a typical row of X in (3), and a scalar y (the corresponding component of the vector y in (3):

(i) Define starting values for the x residuals (e_a) and y residuals (f_a):

e_0= x-□$_x$; f_0= y-□$_y$

For a = 1;2;..., do steps (ii)-(iv) below:

(ii) Introduce scores (t) as linear combinations of the x residuals, using covariates with y as weights (w):

t_a= e_{a-1}'w_a; w_a= Cov(e_{a-1}; f_{a-1}):

(iii) Determine x loadings (p_a) and y loadings (q_a) by least squares: p_a= Cov(e_{a-1}; t_a)=Var(t_a); q_a= Cov(f_{a-1}; t_a)=Var(t_a): (iv) Find new residuals:

ea= e a-1-pa ta; fa= f a-1-qa ta:

This whole procedure, including the construction of scores, loadings and weights,was called 'soft modelling'. At each step r a bilinear representation resulted from this: x= □$_x$+ p_1 t_1+...+ $p_r t_r$+e_r; y = □$_y$+ q_1 t_1+...+q_r t_r+ f_r:

The PLS algorithm is then used for empirical prediction, a method that has become very popular in many applied fields (Mehmood and Ahmed, 2015):

Use the data (X;y), where y now is a vector as in (3). From these data estimate expectations, variances and covariances; otherwise use the procedure to find estimates of scores, loadings and weights. Stop the algorithm at the step r = m, found by crossvalidation. Predict new y-values by using estimates of q_a ,t_a and □$_y$ in the relation y_0= □$_y$+ q_1 t_1+...+q_m t_m.

DOI: https://doi.org/10.5281/zenodo.14513487

All this was met with skepticism by some statisticians; other statisticians chose to neglect the whole development. For those who were engaged, a crucial debate resulted from this; see for instance Martens (1987) and Schweder (1987).

Essentially two problems were the reasons for the statistical skepticism. Firstly, the so-called soft modelling was not related to the 'proper' modelling which statisticians were used to. A model should distinguish sharply between data on the one hand, and parameters on the other hand. Secondly, the algorithm feature and the strongly non-linear characteristic of the procedure, made it very difficult to develop ordinary statistical properties of the method, so that it could be compared to other methods.

The first problem was essentially solved in 1990; the second problem has only begun to reach a solution in the

last few years; see below.

5 The link

In Helland (1990) a proper statistical PLS model was proposed. Take as a point of departure the model (3), where the x-variables are assumed to be random variables, for simplicity centered to zero expectation. Furthermore, we assume that the x-variables have a common covariance matrix $\square_x$ and are uncorrelated with the errors e.

The crucial parameters are then $\square_x$and $\square$. It was shown in Helland (1990) that the scores, loadings and weights in the empirical PLS algorithm could be seen as estimates of scores, loadings and weights in a corresponding population algorithm, made to stop automatically at step m ($e_m = f_m = 0$), (m <min(n; p)). Furthermore, this automatic stop-restriction could be explicitly formulated in two equivalent ways in terms of $\square_x$and $\square$:

(1) Consider the expansion $\square = \square_{i=1}^{p} \square_{id} d_i$, where d_1; ...;$d_p$is a complete set of eigenvectors of$\square\square_x$.

The PLSR model with m factors is defined as the model where this expansion is reduced to exactly m terms. One does not say anything about which eigenvectors are involved in the expansion.

(2) Let $\square = \square_x^{-1}\square\square$, and consider the Krylov sequence $\square;\square_x\square;\square_x^2\square;\square\square_x^3\square\square\square\square\square\square\square$ The PLSR model with m factors is the model where the space spanned by this sequence has dimension m.

These ideas were developed further in a few papers: In Helland (1992) the maximum likelihood estimator of

$\square$under the PLSR model was derived. In Næs and Helland(1993) the concept of relevant components, the eigenvectors correlated with b (cf. formulation (1) above) was discussed in detail, and this concept was related to other regression methods in Helland and Almøy (1994). A group theoretical approach towards the PLSR model was discussed in Helland et al. (2012).

6 Envelope models and PLS

From 2010 on, a new development took place. In the article Cook et al. (2010), the important general envelope model was proposed and discussed, and in Cook et al. (2013)it was shown that a PLSR model - extended to the case of multivariate y - was an important case of the envelope model. For the multivariate case, the regression coefficient $\square\square$in (3) is replaced by a matrix B. The envelope model is - quite similarly to the PLS model, a restriction of the joint parameter ($\square_x$;B). This can be formulated in several ways; one is the following: Assume that there exist an orthogonal matrix

($\square;\square_o$),where $\square$is pxm and $\square_o$is (p-m)xm, and matrices $\square;\square$and $\square_o$ such that B = $\square\square$and

$\square_X = \square\square\square' + \square_o\square_o\square_o'$: (5)

The first term in the expansion for $\square_X$ corresponds to the covariance function for the'relevant' linear combinations of x-variables, the second term to the 'irrelevant' linear combinations. A great number of papers on the envelope model have been published by DennisCook and his collaborators in the last decade, and the theory is summarized in the bookCook (2018). As estimation in this model is concerned, the development is partly dominated by maximum likelihood estimation, which does not work for p >n. A Bayesian approach for envelope models has recently been discussed in Khare et al. (2017).

DOI: https://doi.org/10.5281/zenodo.14513487

What is the situation now? From a statistical point of view, much is still unknown about the PLS algorithm, but see the recent paper by Cook and Forzani (2019) and references there. In the latter paper, asymptotic expansions also in the case when p tends to infinity were studied, and the algorithm was shown to have good statistical properties in many cases when many x-variables have non-zero regression coefficient (no sparsity).

Both for PLS and for the developments initiated by Dennis Cook and his students, it is important that ordinary statistical models can be said to lie behind the predictions done, and the more general applications. A statistical model depends on parameters, and these parameters must be said to have an existence either in the mind of a single researcher, or in the joint minds of a group of researchers. This is a notion that is the background for all statistical modeling, and a notion that we should hold on to, also when trying to approach other areas of (natural) science.

7 Machine learning models

Several statisticians have changed their focus in the last few years. Artificial intelligence, and particularly, machine learning has turned out to be very important in many applications. There are many books on machine learning. A good modern book on the computer approach as seen from statisticians' point of view, is Efron and Hastie (2021).

There is no reason to review the content of the book Efron and Hastie (2021) here, except to emphasize that it is firmly based on statistical models. Chapter 5 discusses models in general, Chapters 8 and 9 uses statistical models as a basis for the authors' treatment of early computer-age methods, while Chapter 16 on sparse modeling is contained in the Section 'Twenty First Century Topics'. But the book also contains a Chapter (18) on neural networks and deep learning.

An important paper on learning rules seen from a Bayesian perspective is Khanand Rue (2021). For those who are interested in the cultural aspect of modern statistical learning, in particular the merging of the model culture and the more algorithmic approach, I mention the brief article Bhadra et al. (2021).

8 Causality and statistical models

The book Pearl (2009) provides a systematic account of cause-effect relationships among variables or events. Modeling causality should be sharply distinguished from statistical modeling. Important notions in Pearl (2009) are directed acyclic graphs, Bayesian networks and the introduction of a do-statement. But this distinction can be formulated in terms of probabilistic modeling, as discussed by Cox and Wermuth (2004). Take as a point of departure three sets of variables B (background), C (causes) and R (response). These can be modeled in recursive form as

$$f_{RCB}=f_{R|CB}f_{B|C}f_{B}. \quad (6)$$

A corresponding full graph of the three sets of variables will contain an arrow from B to R, an arrow from B to C, and an arrow from C to R. With this full graph, the distribution of R, given C can be formed by marginalizing over B:

$$f_{R|C}=\text{integral}\,(f_{R|CB}f_{B|C}db); \quad (7)$$

where $f_{B|C}= f_{BC}/f_{C}$. This is a statistical way of modeling, taking full account of thenotion of conditioning. By contrast, in a causal model the arrow from B to C is assumed to be missing.Then we have a case of intervention from C to R, which can be written as

$$f_{R||C}=\text{integral}\,(f_{R|CB}f_{B}db). \quad (8)$$

The distinction between the two probability distributions $f_{R|C}$and $f_{R||C}$is crucialfor any discussion of inference versus intervention, that is, statistics versus causality.

9 Models for decisions and measurements

In the whole discussion above, it is assumed that the background for modeling, and inparticular the set of statistical parameters, is something that belongs to the mind of a single researcher or to the joint minds of a group of communicating researchers. Let us first assume in general that the mind of some person A contains variables

DOI: https://doi.org/10.5281/zenodo.14513487

□□□□□□□□□□....

In Helland (2021) these are called conceptual variables. Some of the conceptual variables may be decision variables: Let us assume that Ais faced with some concrete decision, having a finite number of prospects $□_1$, $□_2$,..., $□_r$. Then he can define for himself a decision variable, taking the values 1, 2, ... ,r such thatx = k is identified with the choice of the future prospect $□_k$(k = 1, ... , r).One situation lying behind much of the discussion in Helland (2021) is that of a measuring process. A may be planning to do some measurement of a physical variable, say, either □or □or .... These variables may be assumed to be connected to some concrete physical system, but a crucial assumption in Helland (2021) is that they, before, during and after a measurement, also belong to the mind of some person A. This agrees with the philosophy of Convivial Solipsism, as recently was proposed byZwirn (2016) as part of a background for quantum theory: All aspects of the world must be seen from the point of view of the mind of some person. But people can communicate.

Think generally. A is in some specific context, he has his physical environment, his own history, his plans about doing something, and so on. In this context he has several conceptual variables in his mind at some given point of time. Some of these are what I call accessible: It is possible by a measurement or in other ways to find values for them in some future. (Note that this also is meaningful for decision variables.)

We can define a partial ordering both among all conceptual variables and among the accessible ones: Say that □is

'less than or equal to' □if□= f (□) for some function f. I will assume that, if □is accessible, then □= f (□) also is accessible. In many situations A will have in his mind what I call maximally accessible conceptual variables. First look at a measurement situation. Then this may be achieved by variants of Heisenberg's inequality. In particular, position and momentum for a particle at some time will both be maximally accessible. These two variables are then also called, after Niels Bohr, complementary. In the last chapter of Helland (2021), the complementarity notion is attempted generalized to many different areas, also macroscopic.

But go back to the person A and attempt to look at all his conceptual variables at time t from a psychological point of view. We can make a model for this by assuming that there is an inaccessible variable □such that all or a set of accessible variables can be seen as functions of □. In some way □must be said to lie in the subconscious part of A's mind, and it cannot be determined by A himself. A person B, knowing A well and having some insight in practical psychology, may perhaps in some situations come up with a rough estimate of □.

As many models this may be a simplification, but it turns out to be a fruitful simplification.

Relative to this model and using Zorn's Lemma (in the partial ordering □is a dominating conceptual variable), there certainly exist maximally accessible conceptual variables. This is what we need for the next Section.

The paper Yukalov and Sornette (2014) and several related papers lie the foundation of what is called Quantum Decision Theory, a quantum model for decisions made by a single person. In Yukalov et al. (2018, 2022) this theory is extended to joint decisions made by a group of persons. For this, exchange of information is crucial. In Yukalovet al. (2022), collective information shared by the whole group, is also taken into consideration. The decision process for a group of people may take time.

Yukalov and Sornette (2014) derive the quantum probabilities $p(□_i)$, where $□_i$ is the prospect, from what they call a utility factor $f(□_i)$ and an attraction factor $q(□_i)$: $p(□_i)=f(□_i)+q(□_i)$. Here $f(□_i)$ corresponds to the conscious part, modeled through a utility function and a coefficient expressing the actor's belief or confidence.

The factor $q(□_i)$ corresponds to the unconscious part and satisfies $-1 \le q(□_i) \le 1$ and $□_i\ q(□_i) = 0$. In Yukalov et al. (2018, 2022) there is in addition a factor $h(□_i)$ depending on the (joint) information.

DOI: https://doi.org/10.5281/zenodo.14513487

Alternatively, quantum probabilities can be derived from Born's formula.

10 Models as a foundation of quantum mechanics

Several authors have recently tried to rederive the axioms of quantum mechanics, which lead to an abstract theory, from simpler assumptions. As is well known, a basic axiom of quantum mechanics is that a physical system at some fixed time can be associated by a Hilbert space, a complex vector space with a scalar product that is complete in the norm derived from this scalar product, and that the pure states of this system can be defined as unit vectors with arbitrary phase in this Hilbert space. This works well as a basis for quantum calculation, but is very abstract, is difficult to understand for outsiders, and has led to many, partially conflicting, interpretations of the theory.

In Helland (2021, 2022a) this problem is approached by taking conceptual variables as the basic notion. The proofs needed for developing the theory here are somewhat complicated, using group theory and group representation theory, but as we will see, the conceptual basis for the theory can be formulated in a relatively simple way.

The proof of these results relies on two technical theorems. The first one can be formulated as follows:

Theorem 1*Consider a situation where there are two maximally accessible conceptual variables* □*and* □*in the mind of an actor or*
in the minds of a communicating group of actors. Make the following assumptions:

(i) On one of these variables, □*, there can be defined transitive group actions G with a trivial isotropy group and with a left-invariant measure* □*on the space* □□*.*

(ii) There exists a unitary irreducible representation U(.) of the group behind the group actions G defined on □*such that the coherent states U(g)|*□□*>are in one-to-one correspondence with the values of g and hence with the values of* □*.*

(iii) The two maximally accessible variables □*and* □*can both be seen as functions of an underlying inaccessible variable* □*in*□□*. There is a transformation k acting on*□□*such that* □(□□= □(*k*□)*.*

Then there exists a Hilbert space H connected to the situation, and to every accessible conceptual variable there can be associated a symmetric operator on H.

For those that are unfamiliar with the ordinary textbook-formulation of quantum theory, it must be repeated that one of the postulates stated in the textbooks is just that there in every physical situation exists a Hilbert space, and furthermore that every observable variable is associated with a so-called self-adjoint operator in this Hilbert space. Now self-adjoint is in most cases essentially the same as symmetric (Hall, 2013).

The point of the theorem is to derive this conceptual apparatus from simpler assumptions. First look at a measurement situation, as exemplified by measuring either position or momentum of a particle. As discussed in the previous section, these variables can also be seen as conceptual variables, existing in the mind of some observer. For simplicity, look at the one-dimensional case, let □be position, and □be momentum. These are both maximally accessible by the Heisenberg inequality. Let G be the translation group acting on □. Then (i) is satisfied, and (ii) can be shown to be satisfied. The variable f may be the vector (□;□), and based on this vector, the transformation kmay be defined in terms of a Fourier transform from(□;□) to(F[□];F^{-1}[□]). (F is the relevant Fourier transform on □□.

The same theorem can also be used as a basis for a quantum description of a decision situation, if we accept the model described in the previous Section. We then must accept the technical assumptions (i)-(iii), but we are helped by the fact that there already exists a Quantum Decision Theory, founded by Yukalov and Sornette (2014) and related papers.

The second basic theorem, Theorem 4.2 in Helland (2021) is corrected and made more precise in Helland (2022b). To formulate this, we need a definition:

Definition 2*Let* □*(.) be a function from a space* □□*to a space* □□*, and let K bea group of transformations acting on* □□*. The function is said to be permissible (with respect to K) if*

DOI: https://doi.org/10.5281/zenodo.14513487

□(□₁) = □(□□₂) *implies* □(k□₁) = □(k□₂) (9)
for all k ink.

A main result (Helland, 2010) is the following: If □(.) is permissible, then one can define a group action G on
□□and a mapping from K to G such that g(□(□)) = □(k□)(k inK). The mapping from K to G is a homomorphism. The second theorem is now given by

Theorem 2 *Assume that the function □(.) is permissible with respect to the group K. Then, if A□□is the operator corresponding to*
the maximally accessible variable □, we have: For any transformation k in K and any unitary representation V of K, the operator V(k)A□V(k)†is the operator corresponding to □'defined by □'(□) = □(k□).

From these two theorems a rather rich theory follows, as shown in Helland (2021).The theory is general. Operators corresponding to accessible conceptual variables that are not maximally accessible, can be defined by the spectral theorem (Hall, 2013), by taking as a point of departure that if □is any accessible variable, then there exists a maximally accessible □and a function fsuch that □= f(□).The textbook treatment of quantum mechanics concentrates usually on the case where the space □□is finite. For this case it follows (Helland, 2021) from the above theorems: The operator A□□has a finite spectrum. The set of eigenvalues of A□□coincides with the set of possible values for □. The eigenspaces of operators corresponding to a set of conceptual variables □are in one-to-one correspondence with a question 'What will □□be if I measure it?' together with a sharp answer '□=u'. Finally, □is maximally accessible if and only if all eigen spaces of A□□are one-dimensional.

The assumptions of Theorem 1 can tentatively be made more explicit in this finite case. To be precise, the assumption (ii) of Theorem 1 seems to follow in this case (Burnside, 1955) if one can prove that the actual representation V satisfies the following: Every irreducible representation of G occurs as a sub-representation of V^n for sufficiently large n. Ingeneral, we need the existence of a group G and the corresponding representation V={U(.)}Satisfying the assumptions (ii) and (iii). When the dimension of □is 2p for some p, we can use the group corresponding to p independent qubits, and the necessary results are known (Höhn, 2017; Höhn and Wever, 2017).

I will not go into these technicalities here but conclude: The textbook treatment of quantum mechanics can be simplified if one starts with the notion of conceptual variables.

The operator corresponding to an accessible variable can be explicitly defined (Helland, 2021). The eigenvalues of these operators have a simple interpretation. Very many unit vectors in the Hilbert space also have a simple interpretation: They can be interpreted in terms of a question-and-answer pair. To characterize situations where this interpretation is valid for all unit vector in the Hilbert space, is a difficult problem, stated as a question to the quantum community in Helland (2019).

To do calculations, we must express the above in a physical language. Let □□be a maximally accessible conceptual variable taking a finite number of values{u_i; i= 1,...,d}. Then the corresponding Hilbert space is d-dimensional, and can be taken as C^d, the vector space of d-dimensional complex vectors. Following Dirac, the (coloumn) vectors in this space are denoted as |□>, called ket vectors. The corresponding complex conjugate row vectors are denoted by <□| and are called bra vectors. The scalar product can then be defined as <□□|□□>.

Let A□be the operator (matrix) corresponding to □, denote the eigenvectors of A□as {|□ᵢ>; i=1,...,d} and the corresponding eigenvalues by {u_i; i=1,...,d}. Then{|□ᵢ>}is an orthonormal basis of the Hilbert space C^d, and each state vector |□ᵢ>is associated with the event □= u_i.

The state vector|□ᵢ>may equally well be expressed by the projection operator P_i=|□ᵢ><□ᵢ|. The P_i's are orthogonal, and satisfy □ᵢP_i= I, the identity. Let now □be another accessible conceptual variable, not necessarily maximal. Then □is associated with an operator A□, which may have degenerate eigen

DOI: https://doi.org/10.5281/zenodo.14513487

values v_i(i= 1,…,r), where r<d. And again, the eigenspace of A□□corresponding to the eigen value vi can be associated with the statement □= v_i. Let Q_i be the projection upon this eigenspace. Then again, the different Q_i are orthogonal, and □$_i$$Q_i$= I. Each Q_i can, in several ways, be written as a sum of one-dimensional projections on orthogonal eigenvectors.

In the quantum mechanical literature, a ket vector |□>is in general called a state vector, and it characterizes the state of a physical system. In my notation, this usually corresponds to a statement of the type □= u, where □is maximally accessible. The observer A, being in this state relative to the physical system, is then certain of the value of □. Note that □can be a vector, whose components are independent, accessible conceptual variables.

In other cases, A may be uncertain of the value of □□ A special case of this again, is that he for each i has a probability p_i for each value □= u_i. This is also by physicists considered to be a state, a mixed state □. Explicitly

□ =□□$_i$ $p_i$$P_i$, (10)

where P_i is the projector onto the eigenspace of A□□corresponding to □= u_i.

Several possibilities exist for the probabilities p_i. Depending upon the situation, they can be Bayesian priors or posteriors, or they can also have a frequentist interpretation, being derived from a fiducial / confidence distribution. Other possibilities may exist.

It can be verified: □is positive definite and has trace 1. In general, a density operator, characterizing a mixed state is defined as any operator with these two properties.

The probabilities in quantum mechanics are all derived from Born's formula, which canbe derived in several different ways from various conditions and formulated in different ways. One simple variant of such conditions was argued for in Chapter5 of Helland (2021).

11 Discussion

As a general point, it is said in Helland (2021) that the scientific development, both within statistical inference and within quantum physics would perhaps have been slower if a connection between the two cultures had been considered from the beginning.

Universality and creativity can in some sense be seen as complementary properties. At the same way, the complementarity notion may perhaps be used to describe the two scientific societies themselves, but as I have tried to emphasize here, this complementarity is not complete. Links can be found. What can statisticians learn by considering these links? Look first at the PLS modeling case. Here it seems to be clear that the PLS algorithm has a role to play in statistical predictions from regression models with near collinear data. More specifically (Cook and Forzani, 2019) the algorithm seems to give good predictions in cases where, in contrast to the sparsity case, several combinations of x-variables play together in their effect on the response.

Concerning the link between quantum theory and macroscopic activities like statistical inference, it may perhaps be too early to say much about what we can learn. Again, some discussion is given in Helland (2021). It may be of some value for statisticians to look at the notion of taking decisions in the way that is described above. A problem area where this may be investigated, is in connection to the multiplicity issue in relation to the interpretation of p-values, an area where there has been much controversy recently (Wasserstein et al., 2019; Benjami et al., 2021).

In any case, it may be of some value to contrast the use of mathematical modeling in different areas of science.

References

Benjamini, Y., De Veaux, R. D., Efron, B., Evans, S., Glickman, M., Graudbard, B. I., & Kafadar, K. (2021). The ASA president's task force statement on statistical significance and replicability. Annals of Applied Statistics, 15(3), 1084-1085.

DOI: https://doi.org/10.5281/zenodo.14513487

Bhadra, A., Datta, J., Polson, N., Sokolov, V., & Xu, J. (2021). Merging two cultures: Deep and statistical learning. arXiv: 2110.11561 [stat.ME].

Breiman, L. (2001). Statistical modeling: The two cultures. Statistical Science, 16, 199-231.

Burnside, W. (1955). Theory of groups of finite order. Dover Publications, Inc., New York.

Cook, R. D. (2018). An Introduction to Envelopes. Wiley, Hoboken, N.J.

Cook, R. D., Li, B., & Chiaromonte, F. (2010). Envelope models for parsimonious and efficient multivariate linear regression. Statistica Sinica, 20, 927-1010.

Cook, R. D., Helland, I. S., & Su, Z. (2013). Envelopes and partial least squares regression. Journal of the Royal Statistical Society B, 75, 851-877.

Cook, R. D., & Forzani, L. (2019). Partial least squares prediction in high-dimensional regression. Annals of Statistics, 47(2), 884-908.

Cox, D. R., & Wermuth, N. (2004). Causality: A statistical view. International Statistical Review, 72(3), 285-305.

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

Fisher, R. A. (1973). Statistical methods and scientific inference. Hafner Press, New York.

Fraser, D. A. S. (1964). On the definition of fiducial probability. Bulletin of the International Statistical Institute, 40, 842-856.

Hall, B. C. (2013). Quantum Theory for Mathematicians. Springer, New York.

Helland, I. S. (1990). Partial least squares and statistical models. Scandinavian Journal of Statistics, 17, 97-114.

Helland, I. S. (1992). Maximum likelihood regression on relevant components. Journal of the Royal Statistical Society B, 54(2), 637-647.

Helland, I. S. (2010). Steps Towards a Unified Basis for Scientific Models and Methods. World Scientific, Singapore.

Helland, I. S. (2019). When is a set of questions to nature together with sharp answers to those questions in one-to-one correspondence with a set of quantum states? arXiv: 1909.08834 [quant-ph].

Helland, I. S. (2021). Epistemic Processes. A Basis for Statistics and Quantum Theory. Springer, Cham, Switzerland.

Helland, I. S. (2022a). On reconstructing parts of quantum theory from two related maximal conceptual variables. International Journal of Theoretical Physics, 61, 69.

DOI: https://doi.org/10.5281/zenodo.14513487

Helland, I. S. (2022b). The Bell experiment and the limitation of actors. Foundations of Physics, 52, 55.

Helland, I. S., & Almøy, T. (1994). Comparison of prediction methods when only a few components are relevant. Journal of the American Statistical Association, 89(426), 583-591.

Helland, I. S., Sæbø, S., & Tjelmeland, H. (2012). Near optimal prediction from relevant components. Scandinavian Journal of Statistics, 39, 695-713.

Höhn, P. A. (2017). Quantum theory from rules on information acquisition. Entropy, 19(3), 98.

Höhn, P. A., & Wever, C. S. P. (2017). Quantum theory from questions. Physics Reviews A, 95, 012102.

Jöreskog, K. G., & Wold, H. (1982). Systems under indirect observation. Causality structure-prediction. Part I-II. North-Holland, Amsterdam.

Khan, M. E., & Rue, H. (2021). The Bayesian learning rule. arXiv: 2107.04562 [stat.ML].

Khare, K., Pal, S., & Su, Z. (2017). A Bayesian approach to envelope models. Annals of Statistics, 45(1), 196-222.

Martens, H. (1987). Why users like PLS regression. In: Martens, M. (ed.). Data approximation by PLS methods. Report no. 800, Norwegian Computing Centre.

Mehmood, T., & Ahmed, B. (2015). The diversity in the applications of partial least squares: an overview. Journal of Chemometrics, 30(1), 4-17.

Næs, T., & Helland, I. S. (1993). Relevant components in regression. Scandinavian Journal of Statistics, 20, 239-250.

Pearl, J. (2009). Causality. Models, Reasoning, and Inference. 2nd edition. Cambridge University Press, Cambridge.

Pothos, E. M., & Busemeyer, J. R. (2022). Quantum cognition. Annual Reviews of Psychology, 73, 749-778.

Schweder, T. (1987). Canonical regression versus PLS. In: Martens, M. (ed.). Data approximation versus PLS methods. Report no. 800, Norwegian Computing Centre.

Schweder, T., & Hjort, N. L. (2016). Confidence, Likelihood, Probability, Statistical Inference with Confidence Distributions. Cambridge University Press, Cambridge.

Taraldsen, G., & Lindqvist, B. H. (2021). Fiducial inference and decision theory. arXiv: 2112.07060 [stat.ME].

Yukalov, V. I., & Sornette, D. (2014). How brains make decisions. Springer Proceedings Physics, 150, 37-53.

Yukalov, V. I., Yukalova, E. P., & Sornette, D. (2018). Information processing by networks of quantum decision makers. Physics A, 482, 747-766.

DOI: https://doi.org/10.5281/zenodo.14513487