Reply to the discussion of
Invariance and factorial models
by
Peter McCullagh
November 1999

I thank the discussants for their comments and criticisms. Although some might call it abstract nonsense, the theory of categories and functors is nothing more than a mathematical theory of common sense. To say that categories and functors are relevant in applied work is to assert that mathematics and common sense are relevant. While there may be areas of applied work in which neither mathematics nor common sense is a prerequisite, I confine my attention to statistical modelling, an activity that requires both. As Wit notes, the statistical art lies in the construction of one or more categories of relevant morphisms. In an attempt to address some of the points raised, I begin by describing the category structure that is frequently appropriate in the various components of a statistical model.

All sets that arise in statistical work have structure. The purpose of a category is the codification of structure, so category implies structure. Each parametric statistical model considered in this paper has the following components, each with its own structure.

(i) The covariate space is a category in which each object Ω is a set, or product set, of factor levels. The category contains all structure-preserving maps, such as $\mathcal{F}$, $\mathcal{F}^k$ or order-preserving maps.
(ii) The space of units is a category of injective maps such as $\mathcal{I}$ or $\mathcal{ND}$. Each object $\mathcal{U}$ is a set of units; each map $\varphi{:}\,\mathcal{U} \to \mathcal{U}'$ is an injective morphism of units.
(iii) To each unit there corresponds a response or event. The response scale is a category in which each object Γ is a set of mutually exclusive and exhaustive events. The structure of the response scale is determined by the morphisms, which are usually surjective.

Each of these spaces is viewed as a family of similar sets, whose structure is determined by the category of morphisms. Although the distinction between treatment and classification factors is important, I take the view that all specific factors, treatment and classification factors, belong to part (i). Other non-specific factors, primarily relationships among the units such as spatial or temporal relationships or block structure, determine the relevent category of morphisms in part (ii). If the response scale in part (iii) is the real line or a vector space, the morphisms are usually linear or affine; in a multinomial-response model, however, response levels may be amalgamated, leading to non-invertible surjective morphisms.

The design map $\psi{:}\,\mathcal{U} \to \Omega$ associates with each unit, a point in covariate space. It is these maps that Tjur takes as the objects in his category of one-way layouts, and, by extension, to a category of factorial designs. This construction is exactly what is needed to study design-related matters such as estimability, efficiency, computation and so on. I have chosen the simpler construction because the logical properties of models are unaffected by the design map.

In connection with the covariate space, the term factorial structure implies a product category such as $\mathcal{I} \times \mathcal{F}$, $\mathcal{I}^k$ or $\mathcal{F}^k$. Unless the circumstances are exceptional, each component category includes all insertion maps corresponding to subset selection. In most circumstances, other morphisms such as permutations, re-labellings, order-preserving injections, or certain surjections must also be included. Thus, in a factorial structure, each covariate domain is a product set $\Omega = \Omega_1 \times \cdots \times \Omega_k$, and the parameter space $\mathcal{X}_\Omega$ is a *vector subspace $\mathcal{X}_\Omega \subset \mathcal{R}^{\Omega_1\times\cdots\times\Omega_k}$ of certain functions on the levels. On the assumption that the morphisms include permutation and subset selection independently for each factor, it is shown that $\mathcal{X}$ must be a factorial model.*

For two non-homologous factors, the infamous type III sequence,

$$III_{mn} = \{f \in \mathcal{R}^{m\times n} : f_{1\bullet} = \cdots = f_{m\bullet}\}$$

is not closed under selection as the following example demonstrates.

$$\begin{pmatrix} 3 & 7 & 9 & 2 \\ 9 & 5 & 2 & 5 \\ 6 & 2 & 7 & 6 \end{pmatrix} \in III_{34} \quad \xrightarrow{\varphi^*} \quad \begin{pmatrix} 3 & 7 & 2 \\ 9 & 5 & 5 \\ 6 & 2 & 6 \end{pmatrix} \notin III_{33}.$$

As Dawid points out, little knowledge of morphism theory is required to unmask a silly model. *If subset selection is a relevant morphism, as it undoubtedly is in the great majority of applications, any analysis based on such a logical absurdity must be dismissed as nonsense, except insofar as it gives the correct answer by accident, for example in linear models for balanced designs.*

Christensen's calculations demonstrate that the sequence of additive subspaces $A+B \subset A\otimes B$ is selection invariant, but the sequence of 'orthogonal' complements is not. The crux of the matter is the design map $\psi\colon\mathcal{U} \to \Omega_1\times\Omega_2$, and the pullback map $\psi^*\colon A\otimes B \to \mathcal{R}^{\mathcal{U}}$. Two vectors x_1, x_2 that are 'orthogonal' in $A\otimes B$ are carried to two vectors ψ^*x_1, ψ^*x_2, that need not be orthogonal in $\mathcal{R}^{\mathcal{U}}$. The two specifications are thus not equivalent unless the design is complete and balanced. The calculations confirm the assertion in the paper that no sub-representation exists in $A\otimes B$ that is complementary to $A+B$. In so doing, Christensen makes a convincing case for an eleventh commandment: *Thou shalt not use orthogonal complements as models.*

Functional marginality

Suppose that factor A has quantitative levels, so that each finite set $\Omega \subset \mathcal{R}$ is a potential set of levels. The latter condition frequently requires a preliminary logarithmic or other suitable transformation of dose levels. The most natural category of morphisms is frequently the group of affine transformations acting on $\mathcal{R}$, and the corresponding category acting on all subsets by restriction. Then each sub-representation $\mathcal{X}_k$ in $\mathcal{R}^\Omega$ is a vector space consisting of the non-homogeneous polynomials of degree $\le k$. Note that $\mathcal{X}_{k-1}$ is a sub-representation in $\mathcal{X}_k$, but there is no complementary sub-representation. This is Nelder's functional marginality: if x^k is included as a covariate, all lower degree terms must also be included. On the other hand, if zero is a critical point on the scale, the relevant morphisms should preserve this point, possibly by restriction to the linear sub-group. The homogeneous polynomials of degree k constitute an irreducible representation of this sub-group.

Response factor and dual representation

If the response is qualitative, the response scale is a collection of certain finite sets $\Gamma, \Gamma', \ldots$. To say that the response scale is nominal is to say that the category of morphisms includes all surjections; to say that the response scale is ordinal is to say that the category preserves ordinal relationships by including only order-preserving surjections.

Each object $\mathcal{R}^\Gamma$ in the regular contravariant representation is a vector space of *functions* on Γ; in the dual covariant representation, each object is a vector space of *signed* *measures* on Γ. In connection with queries by Forster and Glonek, consider the map $\varphi: \{1,2,3\} \to \{1,2\}$ given by $\varphi(1) = \varphi(2) = 1$; $\varphi(3) = 2$. In the regular contravariant representation, the pullback map $\varphi^*: \mathcal{R}^2 \to \mathcal{R}^3$ is given in matrix form by

$$\varphi^* = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The dual, or adjoint, linear transformation carries the measure P with density (p_1, p_2, p_3) on $\{1,2,3\}$ to the measure $\varphi^{*\prime} P$ with density

$$(\varphi^*)' p = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \begin{pmatrix} p_1 + p_2 \\ p_3 \end{pmatrix}$$

on $\{1,2\}$. In other words, in the dual representation, levels are combined or amalgamated, and probabilities are added in the appropriate manner.

Contingency tables

In the context of contingency tables, each response is a measure on the set of response levels Γ. A log-linear model associates with each point $\eta \in \mathcal{R}^\Gamma/1$ a probability distribution P_η such that

$$P_\eta(\{j\}) = \frac{\exp(\eta_j)}{\sum_{r\in\Gamma} \exp(\eta_r)}.$$

Although P is a vector in the dual space, the parameter η is a point in the space $\mathcal{R}^\Gamma/1$. In a log-linear model for the study of dependence, to each point ω in the covariate space there corresponds a probability distribution $P_{\eta(\omega)}$ on Γ. In this way, a distribution on Γ^Ω is determined by a function $\eta: \Omega \to (\mathcal{R}^\Gamma/1)$ with values in $\mathcal{R}^\Gamma/1$ rather than $\mathcal{R}$. The full parameter space is thus the set of $(\mathcal{R}^\Gamma/1)$-valued functions on Ω, i.e. the representation $(\mathcal{R}^\Gamma/1)^\Omega$, which is isomorphic with $(\mathcal{R}^\Gamma/1) \otimes \mathcal{R}^\Omega$ and with $\mathcal{R}^{\Gamma\times\Omega}/\mathcal{R}^\Omega$. Each subrepresentation, i.e. each factorial model containing $\mathcal{R}^\Omega$, is a log-linear model.

A normal-theory model associates with each point $\omega \in \Omega$ a distribution determined by the parameter $\eta = (\mu, \sigma^2)$ in $\mathcal{V} = \mathcal{R}^2$. In other words, η is a $\mathcal{V}$-valued function on Ω. The full parameter space is thus $\mathcal{V}^\Omega \cong \mathcal{R}^\Omega \oplus \mathcal{R}^\Omega$. Each linear model with constant variance is thus a sub-representation in $\mathcal{R}^\Omega \oplus \mathcal{R}$. A normal-theory 'generalized' linear model is obtained by specifying $g \circ \eta \in \mathcal{X}$ for some suitable link function $g: \mathcal{V} \to \mathcal{V}$. These are the natural transformations referred to by Wit. For various reasons, the statistically natural choice is usually $g = (1, \log)$, leading to linear models for the mean, and multiplicative models for the variance.

Natural manifolds and non-linear models

In connection with Bailey's principle (2), I agree that closure under intersection is an absolute necessity. Although all of the models described in this paper are determined by vector spaces, condition (2) must not be interpreted narrowly in terms of linear models, or even generalized linear models. The preferred version allows not only representations, but also any manifold that is the natural image of a representation. See Wit's definition *of a natural transformation. To give an example, the $\mathcal{F}$-representation $\mathcal{R} \oplus \mathcal{R}^{\Omega}$ may be mapped into $\mathcal{R}^{\Omega^2}$ by the natural transformation*

$$g_{ij}(\theta, \alpha) = \alpha_i \cos\theta + \alpha_j \sin\theta,$$

so the image is the union of the representations $A_\theta \subset A^{\otimes 2}$. For each fixed Ω, the image is a manifold in $\mathcal{R}^{\Omega^2}$, but not a vector subspace. Likewise, for $\mathcal{F}$-representations, the natural transformation $g\colon A^{\oplus 3} \to A^{\otimes 2}$ given by

$$g_{ij}(\alpha, \beta, \gamma) = \alpha_i + \beta_j + \gamma_i\gamma_j$$

has an image that is a smooth sub-manifold within the sub-representation quasi-symmetry. The partial answer to Falgeurolles's question is that this model is an functor on $\mathcal{F}$, and is thus closed under selection and replication, as ordinarily required for application to contingency tables. However, Becker's (1990) model (2) contains a diagonal term, and seems to be closed under selection only. In summary, if non-linear models are permitted, one must allow any natural transformation from a given representation into another representation, provided that the image functor is a sequence of smooth manifolds whose tangent spaces are homomorphic with the given representation.

Homologous factors

The paper states that two factors are homologous if they have the same set of levels. However, the discussion in section 10 makes it clear that the observed sets need not be identical, so the definition is intended to cover Firth's example in which there is a natural map from one set to the other. A more accurate, but less informative, definition is that the two factors are defined as functors on the same category $\mathcal{C}$. Thus, the answer to Preece's first question is that these are two homologous quantitative factors for which each finite subset of $[0, \infty)$ is a set of levels. To say that two factors are homologous is to say that the effect of factor A at level i may be compared with the effect of factor B at the same level. It does not mean that the effects are the same; in competition models, the effects are equal but opposite. Thus, the response function for level r applied in the autumn and level s in the spring might plausibly have the form $0.6\alpha_r + \alpha_s$, which can be justified only if the factors are defined on the same set of levels. As Haberman notes, the homology need not be relevant or helpful. In one such example, (Christensen 1996, table 7.4), the factors are genotype of child and genotype of foster mother. However, the set of legitimate models for homologous factors is a superset of those that do not exploit the homology, so a declaration of homology incurs no loss of generality.

The weighted analysis of the poppy data was used initially to simplify comparisons with similar calculations in the Lane and Lawrence paper. In general, a weighted analysis based on sample variances is not recommended.

In Dr White's example, each domain point is a map from features into variables. Redardless of the experimental design, it seems logically unobjectionable to omit a variable, or to associate several features with the same variable. It appears, therefore, that each domain object is a set of features, a set of variables, and all maps from features into variables. To treat the n observed variables as a subset of an arbitrarily large set is to associate the category $\mathcal{I}$ with the variables. Then, regarding the m features initially as a fixed set, each linear model is a sub-representation of $\mathcal{I}$ in $A^{\otimes m}$. The relevant representations are described in section 4.3, regrettably for $m = 2$ only, but the constructions in section 10 are applicable *to all m. If the set of features is regarded as subject to selection, i.e. if the structure of the features is that of $\mathcal{I}$, some further weeding out is required. Let the m factors be denoted by $A_1, \ldots, A_m$, each with n levels. The representation*

$$\mathrm{symadd}\big(A_1.A_2 + A_1.A_3 \cdots + A_1.A_m + \cdots + A_{m-1}.A_m\big)$$

is an example of an acceptable model of dimension n^2.

The hunt for sub-representations

How does one find the sub-representations of a category $\mathcal{C}$ in a given representation $\mathcal{Z}$? I have found it helpful to ask Burt Totaro, who tells me that the algebraic techniques have not been developed to the algorithmic stage of group representations. Nonetheless, an outline may help. The set of invertible maps $\Omega \to \Omega$ in $\mathcal{C}$ is a group, usually related to the symmetric group. The procedure described in section 9 begins by finding all sub-representations of these groups in $\mathcal{Z}$. Those group representations that are not extendable to a category representation are weeded out. The procedure is clumsy and tedious, but it works well when the core group is well understood and sufficiently large; it fails completely if the core group is the identity, as in order-preserving maps. The suggestion by Cameron and Dawid of going directly to infinity is appealing if it helps to identify the representations more directly.

If the set of levels is fixed and the morphisms are bijective, i.e. permutations and relabellings only, we are in effect using group representations. The answer to Bailey's question preceding (3) is that, for such a category, $T^\perp$ is a representation. However, I think that subset selection is nearly always a relevant morphism in applications, even for factors such as sex, so the question has limited relevance. The diagonal/non-diagonal partition of Ω^2 by injective maps is a different matter. In that case, $A^{\otimes 2}$ contains a two-dimensional subspace of trivial representations.

Bailey asks how to decompose a tensor product of category representations. I rephrase her question (3) in the following form. Let $\mathcal{W}$ be a representation of $\mathcal{C}$, and $\mathcal{W}'$ a representation of $\mathcal{C}'$. How are the sub-representations of the product category in $\mathcal{W} \otimes \mathcal{W}'$ related to the sub-representations in $\mathcal{W}$ and $\mathcal{W}'$? The answer is as follows. To each sub-representation $\mathcal{X} \subset \mathcal{W}$ and $\mathcal{X}' \subset \mathcal{W}'$ there corresponds a product sub-representation $\mathcal{X} \otimes \mathcal{X}' \subset \mathcal{W} \otimes \mathcal{W}'$. The vector span of any subset of these tensor products is a product representation, and thus an allowable model. For reasons that are familiar from group representation theory, this recipe is incomplete if there are multiplicities in $\mathcal{W}$ and $\mathcal{W}'$.

On account of the nested structure of the sibs in example 1, a complete decomposition of Table 1 uses the $\mathcal{ND}$-representations in section 9.5. As Bailey notes, a simpler decomposition based on section 4.3 is incomplete.

Bailey's question (4), concerning the decomposition of $\mathcal{R}^{\Omega^k}$, is difficult enough for group representations. This is another way of saying I do not know the answer except in a few special cases. *Nonetheless, a few general remarks can be made. The* sym, alt *and* plus *operations in section 10 always give $\mathcal{F}$-sub-representations, but not all sub-representations can be generated in this way. The action of the category $\mathcal{I}$ splits the sets Ω^k into orbits. For $k = 3$, each orbit corresponds to a pattern of distinct indices (r,r,r), (r,r,s), (r,s,r), (s,r,r) or (r,s,t). For general k, the orbits are in 1–1 correspondence with the partitions of the set $\{1,\ldots,k\}$. Under certain circumstances, one might also have morphisms acting on k, possibly the group S_k or some sub-group. It is hard to see what category action* *on k would mean. The orbits under the product action $\mathcal{I} \times S_k$ are in 1–1 correspondence with the partitions of the number k.* In the decomposition of $\mathcal{R}^{\Omega^2}$ by sub-representations in section 4.3, I have not required the vector spaces to be closed under transposition of rows and columns. To do so would eliminate the representations, denoted by A_θ for $\theta \neq \pm\pi/4$, including A and B, which is rarely desirable. Likewise, in example 2, I have not required the vector spaces to be closed under any permutation of the four factors, i.e. grandparents. Certain symmetries are a consequence of Mendelian inheritance, but Mendel's laws are not laws of logic. It is not logically impossible that dormancy, like certain forms of colour blindness, is determined entirely through the male line. Each representation can, and should, be classified by such symmetries. Beyond this, I am reluctant to go unless a compelling case is made. The overriding principle is that the family of acceptable models must incorporate all symmetries that are a consequence of logical structure alone, and no others.

Having found all of the sub-representations in $\mathcal{Z}$, they can be displayed, lattice-style as in Fig. 2. An alternative method of description is to find a covering representation $\mathcal{W}$ such that, to each $\mathcal{X} \subset \mathcal{Z}$ there corresponds a homomorphism $g: \mathcal{W} \to \mathcal{Z}$ such that $g\mathcal{W} = \mathcal{X}$. Thus, the $\mathcal{I}$-representation $1 \oplus A$ covers A; the $\mathcal{I}^2$-representation $\mathcal{W} = 1 \oplus A \oplus B \oplus (A \otimes B)$ covers $A \otimes B$. Each homomorphism $g: \mathcal{W} \to A \otimes B$ is a sequence of linear transformations $g_\Omega: \mathcal{W}_\Omega \to (A \otimes B)_\Omega$ having the form

$$g_\Omega(\mu, \alpha, \beta, \gamma)(i, j) = g_0(\mu, \alpha_i, \beta_j, \gamma_{ij})$$

for $\Omega = \Omega_1 \times \Omega_2$, $\mu \in \mathcal{R}$, $\alpha \in A = \mathcal{R}^{\Omega_1}$, $\beta \in B = \mathcal{R}^{\Omega_2}$ and $\gamma \in \mathcal{R}^{\Omega_1 \times \Omega_2}$, in which $g_0: \mathcal{R}^4 \to \mathcal{R}$ is a linear functional on $\mathcal{R}^4$. In other words, the set of homomorphisms $\hom(\mathcal{W}, A \otimes B)$ is a vector space isomorphic with the space of linear functionals on $\mathcal{R}^4$. All of the information in Fig. 2 is contained in the statement that the $\mathcal{F}$-representation

$$\mathcal{W} = 1 \oplus A \oplus (A/1) \oplus \operatorname{sym}^*(A) \oplus \operatorname{alt}^2(A)$$

covers $A^{\otimes 2}$. Each homomorphism $g: \mathcal{W} \to A^{\otimes 2}$ is a sequence of linear transformations $g_\Omega: \mathcal{W}_\Omega \to \mathcal{R}^{\Omega \times \Omega}$ such that

$$g_\Omega(\alpha, \beta, \beta', \theta, \zeta)(i, j) = g_0(\alpha, \beta_i, \beta_j, \beta'_i - \beta'_j, \theta_{ij}, \zeta_{ij})$$

for some linear functional $g_0: \mathcal{R}^6 \to \mathcal{R}$.

Totaro remarks that the representation theory for categories of finite sets may be richer than that for the symmetric group. This is certainly true. All of the representations discussed here have the property that whenever $\mathcal{X}$ is a sub-representation in $\mathcal{W}$, the quotient $\mathcal{W}/\mathcal{X}$ is also a representation. All $\mathcal{F}$-representations have the property that, for each homomorphism $g: \mathcal{X} \to \mathcal{W}$, the kernel of g is a sub-representation in $\mathcal{X}$. In other words, the set of homomorphisms of $\mathcal{F}$-representations is an Abelian category. However, the corresponding statement for $\mathcal{I}$-representations is false.

Exchangeability

I am grateful to Tjur and Keiding for pointing out the connection with Kolmogorov consistency and Rasch's specific objectivity. To this list I might add infinite exchangeability and partial exchangeability (Aldous 1981; Kallenberg 1989). Although Rasch's construction seems more complicated, and more closely related to repetitive structures and extremal families (Lauritzen, 1988), I agree that all these concepts have the same root, namely the category of injective maps.

The word model has been used in this paper in the sense of parameter set. For two homologous factors, Jensen's construction

$$Y_{ij} = X + U_i + V_j + \epsilon_{ij},$$

in which all terms are normally distributed random variables with certain variances and covariances $\mathrm{cov}(U_i, V_i)$, $\mathrm{cov}(\epsilon_{ij}, \epsilon_{ji})$, *determines a specific family of probability measures on the representation* $A^{\otimes 2}$. *Each measure in this family is infinitely exchangeable, in effect a trivial contravariant* $\mathcal{I}$*-representation. An invariant measure* $P = \{(\mathcal{X}_\Omega, \mathcal{A}_\Omega, P_\Omega)\}$ *is a sequence of measure spaces in which (i)* $\mathcal{X}$ *is an* $\mathcal{I}$*-representation; (ii)* $\mathcal{A}_\Omega$ *is the* σ*-algebra of Borel subsets of* $\mathcal{X}_\Omega$*; (iii)* P_Ω *is a measure on the vector space* $\mathcal{X}_\Omega$*. The following diagram illustrates the construction for* $\mathcal{X} = A^{\otimes 2}$ *as a functor on* $\mathcal{I}$*.*

$$\begin{tikzcd}
\mathcal{I} & A^{\otimes 2} & P \\
\Omega \arrow[d, "\varphi"'] & \mathcal{R}^{\Omega\times\Omega} & P_\Omega \\
\Omega' & \mathcal{R}^{\Omega'\times\Omega'} \arrow[u, "\varphi^*"] & P_{\Omega'} \arrow[u, "\varphi^\dagger"]
\end{tikzcd}$$

The measure P *is* $\mathcal{I}$*-invariant, if, for each* $\varphi\colon \Omega \to \Omega'$ *in* $\mathcal{I}$*, the linear map* $\varphi^*\colon \mathcal{X}_{\Omega'} \to \mathcal{X}_\Omega$ *also carries the measure* $P_{\Omega'}$ *to* P_Ω*. In other words the inverse image map* φ^{*-1} *carries each* P_Ω*-measurable set* $A \subset \mathcal{X}_\Omega$ *to a* $P_{\Omega'}$*-measurable set* $\varphi^{*-1}A \subset \mathcal{X}_{\Omega'}$*, and* $P_\Omega(A) = P_{\Omega'}(\varphi^{*-1}A)$*. The vector span of* P *is a one-dimensional trivial* $\mathcal{I}$*-representation because each pair of parallel arrows* $\varphi, \psi\colon \Omega \to \Omega'$ *in* $\mathcal{I}$ *has the same image* $\varphi^\dagger = \psi^\dagger$ *in the representation* $(P, \dagger)$*. To every exchangeable sequence of random variables there corresponds a category of certain injective maps, a representation of this category, and an invariant measure, which is a trivial representation of the category.*

Each parameter point in Jensen's model identifies a particular $\mathcal{I}$*-invariant measure on the representation* $A^{\otimes 2}$*. Most statistical models involve non-trivial representations, and do not have this rather special form.*

Scales of absurdity

Firth describes a number of statistical models that, from the present algebraic viewpoint, look absurd. One must bear in mind, however, that the degree of absurdity is directly proportional to the relevance of the morphisms. Although the type III model scores highly on the absurdity scale, Cox's (1977) example shows that the value is not quite infinite. A generalized linear model for a binary response that uses the logistic link when the number of subjects is even, and a probit link otherwise, is evidently totally absurd however good the fit. Other than functors on categories, a.k.a. common sense, I know of no statistical principle to avoid such absurdities. In another example, if the response for each unit is a bivariate normal random variable, the relevant morphisms $\mathcal{R}^2 \to \mathcal{R}^2$ in part (iii) may be all affine transformations $\mathcal{R}^2 \to \mathcal{R}^2$, or all component-wise affine transformations. If all affine transformations are deemed relevant, the 'seemingly unrelated regressions' phenomenon (Zellner, 1960), is caused by an absurd model. On the other hand, if a convincing argument is made that only component-wise transformations are relevant, the model is not absurd and the phenomenon is 'real'.

Units of measurement
Every statistical object has structure, whether it be physical dimension, units of measurement, factorial structure, or no structure. Category theory forces us to address such issued directly by asking 'what is the scale of measurement?' and 'what are the morphisms?'. By this means, aspects of the formulation that would otherwise be handled by intuition are made explicit. If dimensional analysis and morphism theory had been a routine part of the curriculum, NASA might have been expected to ask in connection with its recent Mars explorer 'what are the units of thrust?' Likewise, the Brother's customers might have asked 'what are the units of medication?' In both cases the resulting disasters could have been averted by attention to the morphisms.

References

Aldous, D. (1981) Representations for partially exchangeable arrays of random variables. *J. Mult. Analysis* **11**, 581–598.

Becker, M. (1980) Quasi-symmetric models for the analysis of contingency tables.

Christensen, R. (1996) *Plane Answers to Complex Questions.* Springer-Verlag, New York.

Fischer, G.H. and Molenaar, I.W. (1994) *Rasch Models.* Springer-Verlag, New York.

Kallenberg, O. (1989) On the representation theorem for exchangeable arrays. *J. Mult. Analysis* **30**, 137–154.

Lauritzen, S, (1988) *Extremal Families and Systems of Sufficient Statistics..* Springer-Verlag, New York.Lecture Notes in Statistics, vol 49.

Zellner, A. (1962) An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias *J. Am. Statist. Assoc.* **57**, 348–368.