



**The University of Chicago
Department of Statistics**

TECHNICAL REPORT SERIES

**Posterior distributions on certain parameter spaces obtained by using
group theoretic methods adopted from quantum physics**

B. Heller^a, M. Wang^{b*}

^aDepartment of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616

^{b*} Department of Statistics, University of Chicago, 5734 S. University Ave. Chicago, IL 60637

TECHNICAL REPORT NO. 546

April 2004

5734 S. University Avenue
Chicago, IL 60637

Posterior distributions on certain parameter spaces obtained by using group theoretic methods adopted from quantum physics

B. Heller^a, M. Wang^{b,*}

^a*Department of Applied Mathematics, Illinois Institute of Technology, Chicago, IL 60616*

^b*Department of Statistics, University of Chicago, 5734 S. University Ave. Chicago, IL 60637*

Abstract

Three one-parameter probability families; the Poisson, normal and binomial distributions, are put into a group theoretic context in order to obtain Bayes posterior distributions on their respective parameter spaces. The families are constructed according to methods used in quantum mechanics involving coherent states. This context provides a method for obtaining noninformative prior measures.

Key words: Non-informative prior or reference prior, group representations, quantum mechanics, spectral measures, coherent states

1. Introduction

The purpose of this paper is to illustrate, in three instances, the construction of posterior probability distributions on parameter spaces using group theoretic methods. Generalizations to further cases are indicated. Group theoretic methods have a natural place in the field of quantum physics. We attempt to co-opt them for use in statistical inference.

1.1. Background

The use of group theoretic methods in statistical inference has been investigated from many points of view. Lists of references are provided, for example, in Eaton(1989), Barndorff-Nielsen et al.(1982), Kass and Wasserman(1996), and Diaconis(1988). Helland(1999 and 2003a, b) provides a scheme for embedding methods of quantum physics into statistical modeling and inference theory via group representation theory.

In this paper we propose, first of all, to use group representations to generate parametric families of probability distributions. Eaton(1989) describes the process in the following manner. Given a measurable space, say, $(\mathcal{X}, \mathcal{B})$ and a group G acting measurably on it, let P_0 be a fixed probability distribution. Then generate a parametric family of probability distributions by $\{gP_0, g \in G\}$. In the spirit of Eaton(1989), we consider probability families generated by group action. However, instead of group elements acting on probability functions (or density functions) we consider group representation operators acting on a complex-valued quantity. Then the probability function is the modulus squared of that quantity. (See the section relating to “quantum mechanics” in Good and Gaskins(1971).) We consider three examples, the Poisson, normal, and binomial families.

*Corresponding Author.

Having generated a parametric family of probability distributions, we obtain a group invariant prior measure on the parameter space. Posterior distributions are then constructed by methods described in the field of quantum mechanics. See, for example, Busch, Grabowski, and Lahti(1995). The three examples given in this paper illustrate the method which is capable of generalization to other probability families. See, for example, Ali, Antoine, and Gazeau(2000).

Along with Malley and Hornstein(1993), and others such as Helland(2003a, b), we anticipate that procedures adopted from the field of quantum physics will become more widely used for purposes of statistical inference. Probability distributions used in quantum physics are obtained in a different manner than those of classical probability. In Parthasarathy(1992), the difference is explained in the following manner. Suppose that we consider the expectation $\mathbb{E}[Y]$ of a real valued discrete random variable. For example, suppose possible value y_i has probability p_i , for $i = 1, 2, \dots, k$. One can express the expectation in terms of the trace of the product of two $k \times k$ diagonal matrices, S and O :

$$\mathbb{E}[Y] = \sum_{i=1}^k p_i y_i = \text{trace} \left[\begin{pmatrix} p_1 & 0 & 0 & \cdots & 0 \\ 0 & p_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & p_k \end{pmatrix} \begin{pmatrix} y_1 & 0 & 0 & \cdots & 0 \\ 0 & y_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & y_k \end{pmatrix} \right] = \text{trace}(SO).$$

In this case since the two matrices are diagonal, they are commutative. In quantum mechanics, noncommutative matrices (or more generally, linear operators) may be used to construct expectations. Here we adopt the same approach. The distinction is sometimes made that the probability distributions used in quantum physics are “noncommutative” as opposed to the “commutative” ones of classical probability.

We begin by showing how to construct noncommutative probability distributions. From there we go on to generate families of probability distributions, and finally, we construct posterior distributions for the three parameter families named above.

1.2. Noncommutative probability distributions

For a description of probability distributions as constructed in quantum mechanics, see, for example, Malley and Hornstein(1993), Whittle(1992), Parthasarathy(1992).

We conceive of a “random experiment” as having two parts. The “input” status is represented by a linear, bounded, Hermitian, positive, trace-one operator S called a *state* operator. For example, if one were tossing a coin, the bias of the coin would be represented by a state operator; loosely speaking, the *state* of the coin. The measurement process (discrete or continuous) or “output” is represented by a linear self-adjoint operator, O , called an *observable* or outcome operator. So that, if one tossed the coin ten times, the measurement process would be to count the number of heads. These linear operators act in a complex separable Hilbert space \mathcal{H} with inner product (\cdot, \cdot) , which is linear in the second entry and complex conjugate linear in the first entry.

Since the observable operator is self-adjoint, it has a real spectrum. We shall consider cases where the spectrum is either all discrete or all continuous. Although operators in a Hilbert space seem far removed from a probability distribution over possible results of an experiment, the relationship is made in the following manner:

- (i) The set of possible (real) results of measurement is the spectrum of the observable operator O . (So, in the coin tossing experiment, O would have a discrete spectrum: $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$.)
- (ii) The expected value for those results, using state operator S , is given by $\text{trace}(SO)$. See Whittle (1992) and Helland(2003a, b).

In order to obtain a probability distribution, the theory then depends upon the spectral theorem for self-adjoint operators. To each self-adjoint O , we can associate a unique set of projection operators $\{\mathcal{E}(B)\}$, for any real Borel set B such that

$$\mathcal{P}\{\text{result} \in B \text{ when the state operator is } S\} = \text{trace}(S \mathcal{E}(B)).$$

This set of projection operators is called a *spectral measure* or a *projection-valued* (PV) measure associated with the self-adjoint operator O . A rigorous definition of PV measure is given in Section 2.5.

There are certain kinds of state operators that are simple to manipulate. They are the projectors onto one-dimensional subspaces spanned by unit vectors φ in the Hilbert space \mathcal{H} . Since each such projection operator is identified by a unit vector in \mathcal{H} , the unit vector itself is called a *vector state*. In this case, the trace formula becomes simplified to an inner product:

$$\text{trace}(S\mathcal{E}(B)) = (\varphi, \mathcal{E}(B)\varphi),$$

where S is the projector onto the one-dimensional subspace spanned by unit vector φ .

Note that if unit vector φ is multiplied by a scalar ϵ of unit modulus, we obtain the same probability distribution as with the vector φ itself. Thus we distinguish between a single unit vector φ and the equivalence class of unit vectors $\{\epsilon\varphi\}$ of which φ is a representative. Similarly as in quantum mechanics, we use the words *vector state* or just *state* to refer to an *arbitrary* representative of a unit vector equivalence class. Thus since, for complex number ϵ of unit modulus,

$$\mathcal{P}\{\text{result} \in B \text{ when the state is } \varphi\} = \mathcal{P}\{\text{result} \in B \text{ when the state is } \epsilon\varphi\},$$

we take φ and $\epsilon\varphi$ to be the same *state* even though they are not the same vectors.

From now on we reserve the use of the word *state* for vector states as described above. To designate a state which is an operator, as opposed to a vector, we use the phrase *state operator*.

1.3. Discrete probability distributions

Consider the case where the spectrum of O is purely discrete and finite, consisting of eigenvalues $\{y_i\}$. Then the eigenvectors $\{\eta_i\}$ of O form a complete orthonormal basis for the Hilbert space \mathcal{H} . In the infinite case, the Hilbert space is realized as an ℓ^2 space of square summable sequences. When the state is φ , the probability of obtaining result y_i is given by $(\varphi, \mathcal{E}(\{y_i\})\varphi)$, where $\mathcal{E}(\{y_i\})$ is the projection onto the subspace spanned by the eigenvectors of the eigenvalue y_i .

In particular, when the spectrum is *simple*, that is, when there is only one eigenvector η_i for each eigenvalue y_i ,

$$\mathcal{P}\{\text{result} = y_i \text{ when the state is } \varphi\} = |(\varphi, \eta_i)|^2. \tag{1.3.1}$$

In order to present examples, we must first decide where to start. The natural method in the performance of statistical inference is to start with a statistical model (say a parametric family of probability distributions) pertaining to the particular physical properties of a given random experiment. Then, perhaps, one may construct posterior distributions on the parameter space based upon observed results. However, here we attempt to construct prototype families for which the “inverse probability” inference procedures that we illustrate below, can be put in place. The examples we present may perhaps fit into the scheme suggested by Helland (2003b).

Thus instead of starting with a statistical model for a particular situation, we start with an observable self-adjoint operator. As this paper progresses, it will become clear that selections of observables and

families of states stem primarily from the selection of a Lie algebra. In this section, however, we consider an example of a PV measure by starting with a given observable operator in the case where its spectrum is discrete and, in fact, finite.

Example 1.3. Consider an experiment with three possible results $1, 0, -1$. Suppose the observable operator is

$$O = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Note that O is Hermitian. The eigenvalues of A are $1, 0, -1$, and corresponding eigenvectors are

$$\eta_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \eta_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \eta_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Once the measurement is represented by a self-adjoint operator O whose eigenvectors serve as a basis for the Hilbert space, then the probability distribution is determined by the choice of state.

Part(a). Consider the unit vector $\xi = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3i \end{pmatrix}$. Using (1.3.1) we have,

$$\begin{aligned} \mathcal{P}_\xi(\text{result} = 1) &= |(\eta_1, \xi)|^2 = \frac{1}{14}, \\ \mathcal{P}_\xi(\text{result} = 0) &= |(\eta_2, \xi)|^2 = \frac{4}{14}, \\ \mathcal{P}_\xi(\text{result} = -1) &= |(\eta_3, \xi)|^2 = \frac{9}{14}, \\ \text{Expected value} \equiv \langle O \rangle_\xi &= (\xi, O\xi) = \frac{1}{14} - \frac{9}{14} = -\frac{8}{14}. \end{aligned}$$

In trace notation,

$$\langle O \rangle_\xi = \text{trace}(SO) = \text{trace} \left[\frac{1}{14} \begin{pmatrix} 1 & 2 & -3i \\ 2 & 4 & -6i \\ 3i & 6i & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \right],$$

where the one-dimensional projector S is obtained by

$$S = \frac{1}{\sqrt{14}} \begin{pmatrix} 1 \\ 2 \\ 3i \end{pmatrix} \frac{1}{\sqrt{14}} (1 \ 2 \ -3i).$$

Part(b). Consider the unit vector $\psi_0 = \frac{1}{2} \begin{pmatrix} -i \\ \sqrt{2} \\ i \end{pmatrix}$.

The probabilities for results $1, 0, -1$ are $\frac{1}{4}, \frac{2}{4}, \frac{1}{4}$ respectively and $\langle O \rangle_{\psi_0} = (\psi_0, O\psi_0) = 0$. We see here how the choice of the state determines the probability distribution.

Part(c). Suppose a family of states can be parameterized by rotation angles (β, θ) of an experimental apparatus in three-dimensional Euclidean space.

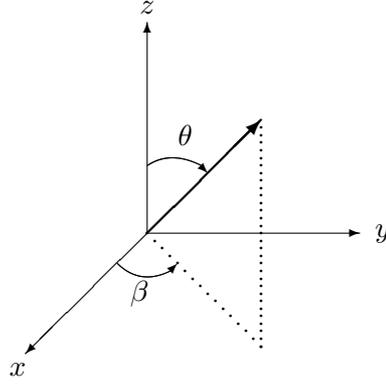


Figure 1: Example 1.3, Part (c).

Suppose the experiment has rotational symmetry in that the probabilistic model does not change when the experimental apparatus is rotated in three-dimensional space. Consider a family of states corresponding to points on the unit sphere indexed by angles β and θ where $0 \leq \beta < 2\pi, 0 \leq \theta < \pi$. Let

$$\psi_{\beta,\theta} = \begin{pmatrix} e^{-i\beta} \cos^2 \frac{\theta}{2} \\ \frac{1}{\sqrt{2}} \sin \theta \\ e^{i\beta} \sin^2 \frac{\theta}{2} \end{pmatrix}.$$

Then,

$$\begin{aligned} \mathcal{P}_{\psi_{\beta,\theta}}(\text{result} = 1) &= |(\eta_1, \psi_{\beta,\theta})|^2 = \cos^4 \frac{\theta}{2}, \\ \mathcal{P}_{\psi_{\beta,\theta}}(\text{result} = 0) &= |(\eta_2, \psi_{\beta,\theta})|^2 = \frac{1}{2} \sin^2 \theta = 2 \sin^2 \frac{\theta}{2} \cos^2 \frac{\theta}{2}, \\ \mathcal{P}_{\psi_{\beta,\theta}}(\text{result} = -1) &= |(\eta_3, \psi_{\beta,\theta})|^2 = \sin^4 \frac{\theta}{2}, \\ \langle O \rangle_{\psi_{\beta,\theta}} &= (\psi_{\beta,\theta}, O\psi_{\beta,\theta}) = \cos^4 \frac{\theta}{2} - \sin^4 \frac{\theta}{2} = \cos \theta. \end{aligned}$$

In Part(b), $\beta = \theta = \frac{\pi}{2}$.

Relabel the possible values:

$$1, 0, -1 \longrightarrow 0, 1, 2,$$

and let

$$p = \sin^2 \frac{\theta}{2}.$$

Then this family becomes the binomial distribution with $n = 2$, and expectation $\langle O \rangle_{\psi_{\beta,\theta}} = 2p$.

It will be seen below that states analogous to $\psi_{\beta,\theta}$ can be obtained for any positive integer n .

1.4. Continuous probability distributions

In the case where the observable self-adjoint operator O has a purely continuous *simple* spectrum $sp(O)$, (that is, there exists a vector ψ_0 in the domain of O such that finite linear combinations of $O^n\psi_0$ are dense in the domain), the Hilbert space is realized as an $\mathcal{L}^2(sp(O), \mu)$ space of complex-valued square integrable functions of a real variable x with inner product

$$(\psi(x), \phi(x)) = \int_{sp(O)} \psi(x)^* \phi(x) \mu(dx),$$

for some finite measure μ with support $sp(O)$, where $*$ indicates complex conjugate. From the spectral theorem (Beltrametti and Cassinelli(1981)), we have the result that self-adjoint operator O determines a unique projection-valued (PV) measure $\{\mathcal{E}(B)\}$ for real Borel sets B . In that case, integrating with respect to the PV measure $\mathcal{E}(B)$, we have formal operator equations:

$$(i) \int_{sp(O)} \mathcal{E}(dx) = I.$$

This should be understood in the sense that $(\psi(x), \mathcal{E}(dx)\psi(x))$ is a probability measure on $sp(O)$, $\forall \psi \in$ the domain of O , since $(\psi(x), \mathcal{E}(B)\psi(x))$ is well defined on any Borel set $B \subseteq sp(O)$.

$$(ii) O = \int_{sp(O)} x \mathcal{E}(dx).$$

It follows that for certain operator functions $f(O)$, we have

$$f(O) = \int_{sp(O)} f(x) \mathcal{E}(dx).$$

In particular, let $\chi_B(x)$ be the characteristic function for Borel set B and let the corresponding operator function be designated as $\chi_B(O)$. Then

$$\chi_B(O) = \int_{sp(O)} \chi_B(x) \mathcal{E}(dx).$$

For vector ξ in the domain of O ,

$$(\xi, \chi_B(O)\xi) = \int_{sp(O)} \chi_B(x) (\xi(x), \mathcal{E}(dx)\xi(x)).$$

When the Hilbert space is constructed in this manner, we say the particular Hilbert space realization is “diagonal in O ” or is “ O -space”. In that case, the O operator is the “multiplication” operator thus: $O\xi(x) = x\xi(x)$, (which explains why the spectral measure for Borel set B is just the characteristic function of that set).

In the diagonal representation of the Hilbert space, since the projection operators $\{\mathcal{E}(B)\}$ are simply the characteristic function $\chi_B(O)$ for that Borel set, we have a simplified form for the probability formula. For unit vector ψ in the domain of O :

$$\mathcal{P}_\psi(O \text{ result} \in B) = (\psi, \chi_B(O)\psi) = \int_B |\psi(x)|^2 \mu(dx).$$

Note that, in this O -diagonal space, the probability distribution is determined by the choice of state ψ .

It is possible to have spectral measures associated with operators which are not self-adjoint. Normal operators also uniquely determine spectral measures but the spectrum might be complex. Subnormal

operators are associated with spectral measures in which the operators $\mathcal{F}(B)$ for complex Borel set B , are not projection operators but are positive operators. We will be using spectral measures of this sort, called “positive-operator-valued” (POV) measures (ref. Section 2.5), instead of projection-valued (PV) measures when we consider probability distributions on parameter spaces.

Example 1.4. We consider the self-adjoint operator Q where $Q\psi(x) = x\psi(x)$, $\psi \in \mathcal{L}^2(sp(Q), dx) \equiv \mathcal{H}$. The Hilbert space \mathcal{H} is diagonal in Q , which represents the measurement of one-dimensional position in an experiment. The spectrum $sp(Q)$ of Q is the whole real line \mathbb{R} .

We choose a state (function of x),

$$\psi(x) = \frac{e^{-x^2/4\sigma^2}}{(2\pi\sigma^2)^{1/4}}, \quad \text{for } \sigma > 0.$$

Then

$$\mathcal{P}_\psi\{\text{position} \in B\} = (\psi(x), \mathcal{E}(B)\psi(x)) = (\psi(x), \chi_B(Q)\psi(x)) = \int_B |\psi(x)|^2 dx.$$

Thus, the probability density function for the distribution is the modulus squared of $\psi(x)$ which is the normal density function with zero mean and variance $= \sigma^2$.

1.5. Groups and group representations

We will consider groups, transformation groups, Lie groups, Lie algebras, and representations of them by linear operators in various linear spaces.

1.5.1. Group representations

Let G be a group with elements denoted by g . Let T denote a linear operator on a linear space \mathcal{H} . We will consider linear spaces over the complex field. In general, we will take \mathcal{H} to be a complex Hilbert space. If, to every $g \in G$, there is assigned a linear operator $T(g)$ such that,

$$(i) \quad T(g_1 g_2) = T(g_1)T(g_2),$$

$$(ii) \quad T(e) = I,$$

where e is the identity element of G and I is the unit operator on \mathcal{H} , then the assignment $g \rightarrow T(g)$ is called a *linear representation* of G by operators T on \mathcal{H} . Usually, the word *linear* is omitted when referring to linear representations. The dimension of the representation is the dimension of the linear space \mathcal{H} .

A representation is called *projective* if (i) above is replaced by

$$(i') \quad T(g_1 g_2) = \epsilon(g_1, g_2) T(g_1)T(g_2), \quad |\epsilon(g_1, g_2)| = 1.$$

A representation is called *unitary* if T is a unitary operator.

Two representations $T(g)$ and $Q(g)$ on linear spaces \mathcal{H} and \mathcal{K} are said to be *equivalent* if there exists a linear operator V , mapping \mathcal{H} into \mathcal{K} with an inverse V^{-1} such that $Q(g) = VT(g)V^{-1}$.

A subspace \mathcal{H}_1 of the space \mathcal{H} of the representation $T(g)$ is called *invariant* if, for $\psi \in \mathcal{H}_1$, $T(g)\psi \in \mathcal{H}_1$ for all $g \in G$. For every representation there are two *trivial* invariant subspaces, namely the whole space and the null subspace. If a representation $T(g)$ possess only trivial invariant subspaces, it is called *irreducible*.

We shall be concerned with *irreducible, projective, unitary* representations of two particular groups.

1.5.2. Transformation groups

By a transformation of a set Ω , we mean a one-to-one mapping of the set onto itself. Let G be some group. G is a *transformation group* of the set Ω if, with each element g of this group, we can associate a transformation $\omega \rightarrow g\omega$ in Ω , where for any $\omega \in \Omega$,

(i) $(g_1g_2)\omega = g_1(g_2\omega)$ and

(ii) $e\omega = \omega$.

A transformation group G on the set Ω is called *effective* if the only element g for which $g\omega = \omega$ for all $\omega \in \Omega$, is the identity element e of G . An effective group G is called *transitive* on the set Ω , if for any two elements $\omega_1, \omega_2 \in \Omega$, there is some $g \in G$, such that $\omega_2 = g\omega_1$. If G is transitive on a set Ω , then Ω is called a *homogeneous space* for the group G .

For example, the rotation group in three-dimensional Euclidean space is not transitive. A point on a given sphere cannot be changed to a point on a sphere of a different radius by a rotation. However, the unit two-sphere is a homogeneous space for the rotation group.

Let G be a transitive transformation group of a set Ω and let ω_o be some fixed point of this set. Let H be the subgroup of elements of G which leave the point ω_o fixed. H is called the *stationary subgroup* of the point ω_o . Let ω_1 be another point in Ω and let the transformation g carry ω_o into ω_1 . Then transformations of the form ghg^{-1} , $h \in H$ leave the point ω_1 fixed. The stationary subgroups of the two points are *conjugate* to each other.

Take one of the mutually conjugate stationary subgroups H and denote by G/H the space of left cosets of G with respect to H . The set G/H is a homogeneous space for G as a transformation group. There is a one-to-one correspondence between the homogeneous space Ω and the homogeneous space G/H . For example, consider the rotation group in three-dimensional space represented by the group of special orthogonal real 3×3 matrices $SO(3)$. The set of left cosets $SO(3)/SO(2)$ can be put into one-to-one correspondence with the unit two-sphere.

1.5.3. Lie algebras and Lie groups

An abstract Lie algebra \mathcal{G} over the complex or real field is a vector space together with a product $[X, Y]$ such that for all vectors X, Y, Z in \mathcal{G} and a, b in the field,

(i) $[X, Y] = -[Y, X]$,

(ii) $[aX + bY, Z] = a[X, Z] + b[Y, Z]$,

(iii) $[[X, Y], Z] + [[Z, X], Y] + [[Y, Z], X] = 0$.

A representation of an abstract Lie algebra by linear operators on a vector space such as a Hilbert space \mathcal{H} , is an algebra homomorphism in the sense that the representation operators have the same product properties as those of the original abstract algebra. For an associative vector space of linear operators, the product operation $[A, B]$ is the commutation operation $[A, B] = AB - BA$.

We will consider representations of two Lie algebras of dimension three. If the basis elements are linear operators E_1, E_2, E_3 , we can indicate a general element as a linear combination $X = aE_1 + bE_2 + cE_3$. The scalar parameters $\{a, b, c\}$ or a subset of them will then become the parameters of the associated probability distribution.

A Lie group is a topological group which is an analytic manifold. The tangent space to that manifold at the identity of the group is called the Lie algebra of the group. It can be shown that the Lie algebra of

a Lie group is an abstract Lie algebra. In the case of a linear representation of a Lie group, the associated Lie algebra can be computed explicitly by differentiating curves through the identity. On the other hand, a (local) Lie group associated with a given Lie algebra can be computed explicitly by the so-called exponential map. (See, for example, Miller(1972)).

For our purposes, we focus upon the parameter space which a Lie group representation inherits from its Lie algebra representation via the exponential map.

1.6. Families of probability distributions

Let G be a group and $g \rightarrow U(g)$ be an irreducible projective unitary representation of G in the Hilbert space \mathcal{H} . For fixed unit vector ψ_0 , the action of $U(g)$ on ψ_0 is designated by,

$$\psi_0 \rightarrow U(g)\psi_0 = \psi_g. \tag{1.6.1}$$

Since each $U(g)$ is unitary, each ψ_g is a unit vector and so can serve as a state. This method of generating a family leads to states designated as ‘‘coherent states’’ in the quantum mechanics literature. In the three families of probability distributions that we consider in detail, the corresponding families of states are coherent states. Perelomov (1986) and Ali, Antoine, Gazeau(2000) give properties of coherent states along with examples and suggestions for generalizations.

Families of states lead to families of probability distributions. Thus, for self-adjoint operator O in O -diagonal Hilbert space \mathcal{H} ,

$$\mathcal{P}\{O \text{ result in Borel set } B \text{ when the state is } \psi_g\} = (U(g)\psi_0, \mathcal{E}(B)U(g)\psi_0) = \sum |U(g)\psi_0, \eta_i|^2$$

in the discrete case, where η_i is the eigenvector corresponding the i th eigenvalue of O , each eigenspace is one-dimensional, and the sum is over all eigenvalues in B , and

$$\mathcal{P}\{O \text{ result in Borel set } B \text{ when the state is } \psi_g\} = (U(g)\psi_0, \mathcal{E}(B)U(g)\psi_0) = \int_B |U(g)\psi_0(x)|^2 \mu(dx)$$

in the continuous case.

Note that the Hilbert space fulfills a double function. On the one hand, it is an O -space. On the other hand it is the space of an irreducible representation of a group. The connection is made in the following way. Consider the Lie algebra \mathcal{G} associated with group G . We choose a basis for \mathcal{G} so that some function of the basis operators is the very self-adjoint operator O that represents possible outcomes of the measured quantity. Then we construct a concrete realization, O -space, of the Hilbert space \mathcal{H} in which we can depict the collection of coherent states along with the concomitant family of probability distributions. This scheme will become transparent when the three examples given below are described in detail.

Clearly, if one has a particular kind of measurement in mind, then the choice of group is limited. Viewing it from the other point of view, choosing a group initially might lead to new parametric families of probability distributions.

2. The Poisson family

We construct the Poisson family by first constructing a particular family of coherent states of the form (1.6.1) in an ℓ^2 Hilbert space \mathcal{H}_N . The family is indexed by a parameter set which also indexes a homogeneous space for a certain transformation group; namely, the Weyl-Heisenberg group, denoted G_W (Perelomov (1986)). Representation operators $T(g), g \in G_W$, acting on a fixed vector in \mathcal{H}_N as in (1.6.1),

generate the coherent states which, in turn, generate the family of probability distributions which leads to the Poisson distribution. This provides a context in which a Bayes posterior probability distribution on the parameter space given an observed Poisson result may be obtained.

2.1. The Weyl-Heisenberg Lie group and Lie algebra and their representations by linear operators

The group G_W can be described abstractly as a three-parameter group with elements $g(s; x_1, x_2)$, for real parameters s, x_1 and x_2 , where the multiplication law is given by

$$(s; x_1, x_2)(t; y_1, y_2) = \left(s + t + \frac{1}{2}(x_1 y_2 - y_1 x_2); x_1 + y_1, x_2 + y_2 \right).$$

Alternatively, we may consider one real parameter s and one complex parameter α , where

$$\alpha = \frac{1}{\sqrt{2}}(-x_1 + ix_2). \quad (2.1.1)$$

Then,

$$(s; \alpha)(t; \beta) = (s + t + \text{Im}(\alpha\beta^*); \alpha + \beta).$$

A concrete realization of the group is given by real 3×3 matrices which are upper triangular with ones on the diagonal. (See Perelomov (1986).)

The Lie algebra \mathcal{G}_W of the Lie group G_W is a nilpotent three-dimensional algebra. Basis elements can be designated abstractly as e_1, e_2, e_3 , with commutation relations

$$[e_1, e_2] = e_3, \quad [e_1, e_3] = [e_2, e_3] = 0.$$

We consider a linear representation of the algebra with basis given by linear operators E_j , for $j = 1, 2, 3$, which operate in a Hilbert space \mathcal{H} . These operators are such that operators iE_j are self-adjoint with respect to the inner product in \mathcal{H} . That property is necessary in order that from this algebra we may construct group representation operators which are unitary. Concrete representation spaces of this algebra where the iE_j are self-adjoint exist and have been constructed (see Perelomov (1986)). Since the representation is an algebra homomorphism, the linear operators E_j have the same commutation relations as the abstract Lie algebra above.

It will prove to be convenient to consider an alternative basis for the three-dimensional linear space of those operators. Put

$$A = \frac{1}{\sqrt{2}}(E_1 - iE_2), \quad A^\dagger = -\frac{1}{\sqrt{2}}(E_1 + iE_2), \quad I = -iE_3.$$

Note that, due to the fact that the iE_j operators are self-adjoint, A^\dagger is indeed an adjoint operator to A . Although A and A^\dagger are not self-adjoint operators, the operator $N = A^\dagger A$ is self-adjoint. We have

$$[A, A^\dagger] = I, \quad [A, I] = [A^\dagger, I] = 0. \quad (2.1.2)$$

A general element of the Lie algebra of operators described above is given by a linear combination of the basis vectors such as

$$X = isI + \alpha A^\dagger - \alpha^* A.$$

This form of linear combination derives from $X = x_1 E_1 + x_2 E_2 + s E_3$ and (2.1.1).

We may now proceed to obtain a group representation by unitary operators T in the Hilbert space \mathcal{H} . By virtue of the exponential map we have,

$$(s; \alpha) \rightarrow T(s; \alpha) = \exp(X).$$

Since the I operator commutes with A and A^\dagger , we may write

$$T(s; \alpha) = e^{is} D(\alpha)$$

where

$$D(\alpha) = \exp(\alpha A^\dagger - \alpha^* A).$$

It is known that this representation is irreducible (Perelomov (1986)).

2.2. The Hilbert space of the irreducible representation

The linear operators mentioned above act in a Hilbert space which has been designated abstractly as \mathcal{H} . In order to consider concrete formulas for probability distributions, it is necessary to give a concrete form to the Hilbert space. In the case of the Poisson family, the space designated \mathcal{H}_N is realized as an ℓ^2 space of complex-valued square summable sequences with basis consisting of the eigenvectors of the self-adjoint operator N . (See Sattinger and Weaver(1986).)

By the so-called ‘‘ladder’’ method, using (2.1.2), it has been found that N has a simple discrete spectrum of non-negative integers. Thus, by the general theory its eigenvectors, $\{\phi_k\}$, $k = 0, 1, 2, \dots$, form a complete orthonormal set in \mathcal{H} which forms a basis for the ℓ^2 Hilbert space realization \mathcal{H}_N . See, for example, McMurry(1994).

In \mathcal{H}_N , we have the following useful properties of the A (annihilation), A^\dagger (creation), and N (number) operators.

$$\begin{aligned} A\phi_0 &= 0, & A\phi_k &= \sqrt{k} \phi_{k-1} & \text{for } k &= 1, 2, 3, \dots, \\ A^\dagger\phi_k &= \sqrt{k+1} \phi_{k+1} & \text{for } k &= 0, 1, 2, 3, \dots, \\ N\phi_k &= k\phi_k & \text{for } k &= 0, 1, 2, 3, \dots. \end{aligned} \quad (2.2.1)$$

Then we can relate ϕ_k to ϕ_0 by

$$\phi_k = \frac{1}{\sqrt{k!}} (A^\dagger)^k \phi_0 \quad \text{for } k = 0, 1, 2, \dots. \quad (2.2.2)$$

2.3. Family of coherent vectors

To construct a family of coherent vectors in \mathcal{H}_N , the first step is to choose and fix a vector in \mathcal{H}_N . It is convenient to choose the ‘‘ground state’’ basis vector ϕ_0 for that purpose. Then define the family of vectors

$$v(s; \alpha) = T(s; \alpha)\phi_0 = e^{is} D(\alpha)\phi_0. \quad (2.3.1)$$

To find an explicit formula for the v vectors, write $D(\alpha)$ as a product of exponential operators. Since A and A^\dagger do not commute, we do not have the property which pertains to scalar exponential functions that $D(\alpha) = \exp(\alpha A^\dagger) \exp(-\alpha^* A)$. We use the Baker-Campbell-Hausdorff operator identity:

$$\exp(O_1) \exp(O_2) = \exp\left(\frac{1}{2}[O_1, O_2]\right) \exp(O_1 + O_2) \quad (2.3.2)$$

which is valid in the case where the commutator $[O_1, O_2]$ commutes with both operators O_1 and O_2 . Putting $O_1 = \alpha A^\dagger$, $O_2 = -\alpha^* A$, we have

$$D(\alpha) = \exp(\alpha A^\dagger - \alpha^* A) = e^{-|\alpha|^2/2} \exp(\alpha A^\dagger) \exp(-\alpha^* A).$$

Thus,

$$T(s; \alpha)\phi_0 = e^{is}D(\alpha)\phi_0 = e^{is}e^{-|\alpha|^2/2} \exp(\alpha A^\dagger) \exp(-\alpha^* A) \phi_0.$$

For linear operators, we have the same kind of expansion for an exponential operator function as for a scalar function. Expanding $\exp(-\alpha^* A)\phi_0$ and using (2.2.1), we find that $\exp(-\alpha^* A)\phi_0 = I \phi_0$. Then from (2.2.2), we see that

$$(A^\dagger)^k \phi_0 = \sqrt{k!} \phi_k.$$

We have,

$$\exp(\alpha A^\dagger)\phi_0 = \left(I + \alpha A^\dagger + \frac{\alpha^2}{2!}(A^\dagger)^2 + \cdots + \frac{\alpha^k}{k!}(A^\dagger)^k + \cdots \right) \phi_0.$$

From (2.2.1) and (2.3.1),

$$v(s; \alpha) = T(s; \alpha)\phi_0 = e^{is}e^{-|\alpha|^2/2} \sum_{k=0}^{\infty} \frac{\alpha^k}{\sqrt{k!}} \phi_k. \quad (2.3.3)$$

2.4. Family of probability distributions

Let the observable (self-adjoint) number operator N represent the physical quantity being “counted” with possible outcomes $0, 1, 2, \dots$. Using the family of coherent vectors given above, the probability distributions are,

$$\mathcal{P}_{v(s; \alpha)}\{\text{result} = n\} = |(\phi_n, v(s; \alpha))|^2.$$

By expression (2.3.3) for $v(s; \alpha)$, we have the inner product

$$(\phi_n, v(s; \alpha)) = e^{is}e^{-|\alpha|^2/2} \sum_{k=0}^{\infty} \frac{\alpha^k}{\sqrt{k!}} (\phi_n, \phi_k). \quad (2.4.1)$$

Then the orthonormality of the basis vectors, ϕ_k gives

$$(\phi_n, v(s; \alpha)) = e^{is}e^{-|\alpha|^2/2} \frac{\alpha^n}{\sqrt{n!}}. \quad (2.4.2)$$

Taking the modulus squared, we have the formula for the Poisson family,

$$\mathcal{P}_{v(s; \alpha)}\{\text{result} = n\} = e^{-|\alpha|^2} \frac{(|\alpha|^2)^n}{n!} \quad \text{for } n = 0, 1, 2, 3, \dots$$

Put the Poisson parameter $\lambda = |\alpha|^2$. Thus we see that λ is real and nonnegative. Note that the s parameter disappeared when we took the modulus squared of the inner product.

It may be remarked that this is a complicated method for obtaining the Poisson family. The point is that we now have a context in which to infer a probability distribution on the parameter space, given an observed Poisson value n .

2.5. POV measures versus PV measures

Consider the definition of a projection-valued (PV) measure, or *spectral measure* (see, for example, Busch, Grabowski, and Lahti (1995)), which had been introduced heuristically in Section 1.2.

Definition. Let $\mathcal{B}(\mathbb{R})$ denote the Borel sets of the real line \mathbb{R} and $\Lambda(\mathcal{H})$ denote the set of bounded linear operators on \mathcal{H} . A mapping

$$\mathcal{E} : \mathcal{B}(\mathbb{R}) \rightarrow \Lambda(\mathcal{H})$$

is a projection valued (PV) measure, or a spectral measure, if

$$\begin{aligned} \mathcal{E}(B) &= \mathcal{E}^\dagger(B) = \mathcal{E}^2(B), & \forall B \in \mathcal{B}(\mathbb{R}), \\ \mathcal{E}(\mathbb{R}) &= I, \\ \mathcal{E}(\cup_i B_i) &= \sum_i \mathcal{E}(B_i) \quad \text{for all disjoint sequences } \{B_i\} \subset \mathcal{B}(\mathbb{R}), \end{aligned} \tag{2.5.1}$$

where the series converges in the weak operator topology.

This spectral measure gives rise to the definition of a unique self-adjoint operator O defined by

$$O = \int x \mathcal{E}(dx),$$

with its domain of definition

$$\mathcal{D}(O) = \left\{ \psi \in \mathcal{H}, \text{ s.t. } \left(\psi, \int x^2 \mathcal{E}(dx) \psi \right) = \int x^2 (\psi, \mathcal{E}(dx) \psi) \text{ converges} \right\}. \tag{2.5.2}$$

Given a self-adjoint operator O with domain of definition $\mathcal{D}(O) \subseteq \mathcal{H}$, there is a unique PV measure $\mathcal{E} : \mathcal{B}(\mathbb{R}) \rightarrow \Lambda(\mathcal{H})$ such that $\mathcal{D}(O)$ is (2.5.2), and for any $\psi \in \mathcal{D}(O)$,

$$(\psi, O\psi) = \int x (\psi, \mathcal{E}(dx) \psi).$$

Therefore there is a one-to-one correspondence between self-adjoint operators O and real PV measures \mathcal{E} .

In the case of the Poisson distribution, the self-adjoint operator is N with normalized eigenvectors $\{\phi_k\}$ as the orthonormal basis of \mathcal{H} , and $sp(N) = \{0, 1, 2, \dots\}$. Using the ‘‘coherent vectors’’ $v(s; \alpha)$ in the previous section, we obtained the Poisson distribution.

For an inferential probability measure operator on the parameter space associated with the Poisson distribution, we will have neither a PV measure nor a self-adjoint operator. Instead we will have a *subnormal* operator and a so-called positive operator-valued (POV) measure, where the first line of (2.5.1) is amended to read

$$\mathcal{E}(B) \text{ is a positive operator for all } B \in \mathcal{B}(\mathbb{R}).$$

The operators for PV measures are *projections*. The properties prescribed for those projections $\mathcal{E}(B)$ are just those needed so that the corresponding inner products $(\psi, \mathcal{E}(dx) \psi)$ for vector states ψ (or more generally traces with state operators) will have the properties of probabilities. However, for the definition of probability there is no requirement that the operators be projections. In fact, if they are *positive* operators, they can still lead to probabilities when put together with states. That is, for a POV measure \mathcal{F} , $(\psi, \mathcal{F}(\cdot) \psi)$ is still a well defined probability measure on the spectrum of the subnormal operator (Busch, Grabowski, and Lahti(1995)).

2.6. An invariant measure on the parameter space

Now we consider the inferential or so-called *retrodictive* case as described in Peres (1995). In a sense we reverse the roles of states and observable operators. If the Poisson value n was observed, what was formerly a vector ϕ_n denoting a one-dimensional projection $\mathcal{E}(\{n\})$, now becomes a *state*. What was formerly a

family of coherent vectors, $v(s; \alpha)$, now leads to the construction of a POV measure on (a modified version of) the parameter set.

The context for the construction of the Poisson distribution via coherent states, which derives from an irreducible representation of the Weyl-Heisenberg group provides a method for obtaining an invariant measure on the parameter space which, in a Bayes context can serve as a kind of “prior measure”. However, there is a caveat. First it is necessary to be more explicit about the nature of the parameter space.

Recall that the parameters (real s and complex α) first appeared as coefficients in a linear combination of the three basic linear operators $\{A, A^\dagger, I\}$ for the representation of the Weyl-Heisenberg Lie algebra. The subsequent group representation is expressed as

$$g(s; \alpha) \rightarrow T(s; \alpha) = e^{is}D(\alpha).$$

Thus the parameters s and α index the group elements. That implies they index the coherent vectors which implies they index the inner product whose modulus squared becomes the Poisson family.

Recall that the s parameter disappears when we take the modulus squared. We should not have considered it part of the parameter space in the first place. Certainly we do not want to consider it part of the parameter space for an inferential (retrodictive) probability distribution.

Thus we re-examine the definition of the parameter space.

Consider the distinction between a *vector* in \mathcal{H} and a *state* in \mathcal{H} , described in Section 1.2. Specifically, in terms of coherent *states*, if

$$v_1 = T(g_1)\phi_0 \quad \text{and} \quad v_2 = T(g_2)\phi_0 = \epsilon T(g_1)\phi_0 \tag{2.6.1}$$

where $|\epsilon| = 1$, then v_1 and v_2 , although not the same as vectors, indeed are the same as *states*.

The expression (2.6.1) holds if and only if

$$T(g_1^{-1}g_2)\phi_0 = \epsilon \phi_0.$$

Let H_W be a subgroup of G_W such that

$$T(h)\phi_0 = \epsilon \phi_0 \quad \text{for} \quad h \in H_W.$$

When a subgroup such as H_W is maximal, it is called the *isotropy subgroup* for the state ϕ_0 . The set of left cosets G_W/H_W is indexed by parameter $\alpha \in \mathbb{C}$ instead of the parameters $(s; \alpha)$ of the group G_W . This set of left cosets is a homogeneous space for the action of G_W and is isomorphic to the complex plane.

Consider the states $w(\alpha)$ corresponding to the family of coherent vectors $v(s; \alpha) = e^{is}w(\alpha) = e^{is}D(\alpha)\phi_0$. Any $\alpha \in \mathbb{C}$ indexes a coset in G_W/H_W and corresponds to a state $w(\alpha)$. For $g(\alpha_1), g(\alpha_2) \in G_W/H_W$, $T(g(\alpha_1))\phi_0 \neq T(g(\alpha_2))\phi_0$ if g_1 and g_2 are not from the same coset, thus different cosets correspond to different coherent states. Therefore the set of coherent states $\{w(\alpha)\}$ is in one-to-one correspondence with its index set, the complex plane.

In order to obtain a measure on the parameter space \mathbb{C} which is invariant to an operator $D(\beta)$, for arbitrary complex number β , we need to see how the operator transforms a coherent state $w(\alpha)$. Consider

$$D(\beta)w(\alpha) = D(\beta)D(\alpha)\phi_0.$$

Using the Baker-Campbell-Hausdorff identity (2.3.2) with $O_1 = \beta A^\dagger - \beta^* A$ and $O_2 = \alpha A^\dagger - \alpha^* A$, we have

$$D(\beta)D(\alpha)\phi_0 = e^{i\text{Im}(\beta\alpha^*)}D(\beta + \alpha)\phi_0.$$

As states,

$$D(\beta)w(\alpha) = w(\beta + \alpha).$$

Thus, the operator D acts as a translation operator on the complex plane so that the invariant measure, $d\mu(\alpha)$, is just Lebesgue measure, $d\mu(\alpha) = cd\alpha_1d\alpha_2$, where $\alpha = \alpha_1 + i\alpha_2$, and where c is a constant. For normalization purposes, we take $c = 1/\pi$.

2.7. A posterior distribution on the parameter space

We construct a POV measure on the parameter space which leads to a Bayes posterior distribution.

By general group theory, the irreducibility of the group representation by unitary operators in the Hilbert space \mathcal{H} implies that the coherent states are complete in \mathcal{H} . (See, for example, Perelomov(1986)). Thus, for any vectors ψ_1 and ψ_2 in \mathcal{H} , we have,

$$(\psi_1, \psi_2) = \int (\psi_1, w(\alpha)) (w(\alpha), \psi_2) d\mu(\alpha).$$

The coherent states form a so-called ‘‘overcomplete’’ basis for a Hilbert space in the sense that they are complete and can be normalized but are not orthogonal (or even linearly independent). The Hilbert space \mathcal{H}_{CS} which they span may be visualized as a proper subspace of $L^2(\mathbb{C})$, the space of square integrable functions of a complex variable with inner product $(f(\alpha), g(\alpha))_{CS} = \int_{\mathbb{C}} f(\alpha)^* g(\alpha) d\mu(\alpha)$. In Ali, Antoine, Gazeau(2000), an isometric map ρ is given which associates an element ϕ in \mathcal{H}_N with element (function of α) in \mathcal{H}_{CS} :

$$\rho(\phi) = (\phi, w(\alpha)),$$

where the inner product is that of \mathcal{H}_N .

Using (2.3.3), we see that, similarly as in (2.4.1),

$$\rho(\phi) = e^{-|\alpha|^2/2} \sum_{k=0}^{\infty} \frac{\alpha^k}{\sqrt{k!}} (\phi, \phi_k),$$

from which it can be shown that the map is isometric.

In \mathcal{H}_{CS} we construct a POV measure M which leads to a probability distribution for α defined by

$$\mathcal{P}\{\alpha \in \Delta \text{ for state } \psi\} = (\psi, M(\Delta)\psi) = \int_{\Delta} |(\psi, w(\alpha))|^2 d\mu(\alpha)$$

for complex Borel set Δ and for state ψ . (See Busch, Grabowski, and Lahti(1995).)

In particular, consider the (eigenvector) state $\psi = \phi_n$ corresponding to an observed Poisson value n , an eigenvalue of the self-adjoint number operator N .

$$\mathcal{P}\{\alpha \in \Delta \text{ for state } \phi_n\} = \int_{\Delta} |(\phi_n, w(\alpha))|^2 d\mu(\alpha).$$

This provides us with a probability distribution on the whole parameter space, namely, the complex plane. But the Poisson parameter, a real number, is the modulus squared $|\alpha|^2$ of α . Expressing α in polar coordinates, $\alpha = re^{i\theta}$, with $r > 0$ and $0 \leq \theta < 2\pi$, we obtain the invariant measure on the plane,

$$d\mu(\alpha) = \frac{1}{2\pi} dr^2 d\theta = \frac{1}{\pi} r dr d\theta. \quad (2.7.1)$$

Then integrating θ from 0 to 2π , we obtain the marginal distribution for r^2 as follows. For real Borel set B ,

$$\begin{aligned} \mathcal{P}\{r^2 \in B \text{ for state } \phi_n\} &= \int_B e^{-r^2} \frac{(r^2)^n}{n!} \left(\frac{1}{2\pi} \int_0^{2\pi} d\theta \right) dr^2 \\ &= \int_B e^{-\lambda} \frac{\lambda^n}{n!} d\lambda, \end{aligned}$$

where $\lambda = r^2$ and the expression for $|(\phi_n, w(\alpha))|^2$ is obtained similarly as in (2.4.2).

We note parenthetically that the subnormal operator associated with this POV measure is the A operator as shown in Perelomov(1986). As mentioned in Busch, Grabowski, and Lahti(1995), for subnormal operators, the association with a POV measure is not unique. Our choice, in this instance, was motivated by the effort to produce a Bayes posterior distribution. The construction above basically stipulates a Bayes invariant prior measure on the (complex) α parameter space. It is interesting to note that the probability distribution for the (real) λ parameter space only appears in the subsidiary role of a marginal distribution. A decision theoretic optimization method is described in Helstrom (1976) for choosing a POV measure which leads to the POV choice given above.

3. The normal translation family

We construct the normal translation family similarly as the Poisson family. The relevant Lie group is the Weyl-Heisenberg group just as it is for the Poisson family. The only difference is the concrete realization of the Hilbert space. For the Poisson family, the Hilbert space has a countable basis comprised of the eigenvectors of the self-adjoint number operator N . For the normal family, the Hilbert space is the \mathcal{L}^2 space of complex-valued square-integrable functions of a real variable.

3.1. Lie algebra and Lie group representation

In section 2.1, for the Poisson family, we focused on the three basic operators $\{I, A, A^\dagger\}$ for the representation of the Weyl-Heisenberg Lie algebra. For purposes of the normal family, we consider an alternative basis of operators $\{Q, P, I\}$ with

$$Q = \sigma(A + A^\dagger), \quad P = -i\sigma(A - A^\dagger),$$

where σ is a positive real constant. We have the commutation relations

$$[Q, P] = i2\sigma^2 I, \quad [Q, I] = [P, I] = 0.$$

Denoting two real parameters by q and p , a general element of the Lie algebra is given by,

$$X = isI + \frac{i}{2\sigma^2}(pQ - qP).$$

This X corresponds to $X = isI + (\alpha A^\dagger - \alpha^* A)$ of Section 2.1 where q and p are related to complex α by

$$\alpha = \frac{1}{2\sigma}(q + ip). \tag{3.1.1}$$

As in Section 2.1, the group representation operators are obtained from the Lie algebra operators by the exponential map and we have the irreducible group representation by unitary operators:

$$T(s, q, p) = e^{is} D(q, p), \quad \text{where} \quad D(q, p) = \exp \left[\frac{i}{2\sigma^2}(pQ - qP) \right].$$

Thus far, the only difference between this presentation and that for the Poisson distribution is a re-parameterization of the group parameters from $\{s, \alpha\}$ to $\{s, q, p\}$.

3.2. A concrete realization of the Hilbert space of the irreducible representation

Here we have the distinction between the Poisson family and the normal translation family. In the case of the normal family, The Hilbert space takes the form known as Schrödinger space:

$$\mathcal{H}_{sch} = \mathcal{L}^2(\mathbb{R}, dx),$$

the space of complex-valued square integrable functions of real variable x where the Q operator is the “multiplication” operator,

$$Q\psi(x) = x\psi(x).$$

The inner product in \mathcal{H}_{sch} is

$$(\psi(x), \phi(x)) = \int_{\mathbb{R}} \psi(x)^* \phi(x) dx. \quad (3.2.1)$$

The P operator is the differential operator given by $P\psi(x) = -i2\sigma^2\psi'(x)$.

Both Q and P are unbounded operators with continuous spectra. It can be shown that their domains are dense in \mathcal{H}_{sch} and that they are self-adjoint. In quantum mechanics, the operator Q relates to the measurement of one-dimensional position and P to linear momentum.

3.3. Family of coherent states and the normal translation family

As described in Section 1.4, probability distributions relating to self-adjoint operators such as Q with continuous spectra, are given by integrals where the integrand is the modulus-squared of a state function $\psi(x)$. As before, we construct a family of probability distributions by constructing a family of states which in this case are functions of a real variable x .

In the Poisson case, the family of coherent states is of the form $D(\alpha)\phi_0$, where the vector ϕ_0 is the eigenvector of N corresponding to the lowest eigenvalue 0. In the normal case, we have exactly the same construction with the re-parameterization given in (3.1.1) and the eigenvectors of the operator N written as elements (functions) in \mathcal{H}_{sch} .

In \mathcal{H}_{sch} , the set of eigenvectors of N are the functions

$$\phi_n(x) = \frac{1}{(2\pi\sigma^2)^{1/4}(2^n n!)^{1/2}} H_n \left(\frac{x}{\sqrt{2\sigma^2}} \right) e^{-x^2/4\sigma^2}, \quad \text{for } n = 0, 1, 2, \dots,$$

where the functions H_n are the Hermite polynomials. Thus,

$$\phi_0(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-x^2/4\sigma^2}.$$

Then $D(q, p)\phi_0$ in \mathcal{H}_{sch} is the family of coherent state functions,

$$w_{q,p}(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} e^{-ipq/4\sigma^2} e^{ipx/2\sigma^2} e^{-(x-q)^2/4\sigma^2}. \quad (3.3.1)$$

The modulus squared of those functions is the density function for the normal family with fixed value for σ^2 and translation parameter q . Note that when we take the modulus squared, the parameter p drops out.

3.4. An inferred probability distribution on the parameter space

Similarly as for the Poisson distribution, the parameters $\{q, p\}$ index a homogeneous space, the plane, which is isomorphic to the set of left cosets given in Section 2.6 and the action of $D(q, p)$ on that plane is one of translation. Thus, we have the invariant measure $(1/(4\pi\sigma^2))dqdp$ on the parameter space.

Also, similarly as in Section 2.7, the coherent states $w_{q,p}(x)$ are complete in the Hilbert space \mathcal{H}_{sch} . Completeness implies that for vectors ψ_1 and ψ_2 in \mathcal{H}_{sch} , we have,

$$(\psi_1, \psi_2) = \frac{1}{4\pi\sigma^2} \iint_{\mathbb{R}^2} (\psi_1, w_{q,p})(w_{q,p}, \psi_2) dqdp, \quad (3.4.1)$$

where $\frac{1}{4\pi\sigma^2}dqdp$ corresponds to $\frac{1}{\pi}d\alpha_1d\alpha_2$ using (2.7.1) and (3.1.1) with $\alpha = \alpha_1 + i\alpha_2$. Writing the \mathcal{H}_{sch} inner product in (3.4.1) as in (3.2.1),

$$(\psi_1(x), \psi_2(x)) = \frac{1}{4\pi\sigma^2} \iint_{\mathbb{R}^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi_1(x')^* w_{q,p}(x') w_{q,p}(x'')^* \psi_2(x'') dx' dx'' dqdp. \quad (3.4.2)$$

In Ali, Antoine, and Gazeau(2000) it is shown that there is an isometric map from \mathcal{H}_{sch} to the space $\mathcal{H}_{CS}^{q,p}$ spanned by the coherent states.

As in Busch, Grabowski, and Lahti(1995), we construct a POV measure in $\mathcal{H}_{CS}^{q,p}$ which leads to the probability distribution for $\{q, p\}$ in state ψ given by,

$$\mathcal{P}_\psi\{(q, p) \in \Delta\} = \frac{1}{4\pi\sigma^2} \iint_{\Delta} (\psi, w_{q,p})(w_{q,p}, \psi) dqdp. \quad (3.4.3)$$

We proceed to find the marginal distribution of parameter q using (3.3.1), (3.4.2) and (3.4.3).

$$\begin{aligned} \mathcal{P}_\psi\{(q, p) \in \Delta\} &= \frac{1}{4\pi\sigma^2} \iint_{\Delta} \int_{\mathbb{R}} \int_{\mathbb{R}} \psi(x')^* w_{q,p}(x') w_{q,p}(x'')^* \psi(x'') dx' dx'' dqdp \\ &= \frac{1}{4\pi\sigma^2} \iint_{\Delta} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{ip(x'-x'')}{2\sigma^2}} e^{-\frac{(x'-q)^2}{4\sigma^2}} e^{-\frac{(x''-q)^2}{4\sigma^2}} \psi(x')^* \psi(x'') dx' dx'' dqdp. \end{aligned} \quad (3.4.4)$$

For a posterior distribution, we select a state that represents an observed value, i.e. an element of the spectrum of Q , a real number y . Since the spectrum is continuous, y in the form of a ‘‘function’’ of x is represented as a Dirac delta function:

$$y(x) = \delta(y - x).$$

Putting $\delta(y - x)$ for $\psi(x)$ in (3.4.4) we have

$$\frac{1}{4\pi\sigma^2} \iint_{\Delta} \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{ip(x'-x'')}{2\sigma^2}} e^{-\frac{(x'-q)^2}{4\sigma^2}} e^{-\frac{(x''-q)^2}{4\sigma^2}} \delta(x' - y) \delta(y - x'') dx' dx'' dqdp.$$

In order to obtain the marginal distribution of q , we integrate p over the whole real line. For real Borel set B ,

$$\begin{aligned} &\mathcal{P}\{q \in B \text{ when the state is (observed) } y\} \\ &= \int_B \int_{\mathbb{R}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} \delta(x' - y) \delta(y - x'') e^{-\frac{(x'-q)^2}{4\sigma^2}} e^{-\frac{(x''-q)^2}{4\sigma^2}} \left(\frac{1}{4\pi\sigma^2} \int_{\mathbb{R}} e^{\frac{ip(x'-x'')}{2\sigma^2}} dp \right) dx' dx'' dq. \end{aligned}$$

Using

$$\frac{1}{4\pi\sigma^2} \int_{\mathbb{R}} e^{\frac{ip(x'-x'')}{2\sigma^2}} dp = \delta(x' - x''),$$

we have

$$\begin{aligned}
& \mathcal{P}\{q \in B \text{ when the state is (observed) } y\} \\
&= \int_B \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} \delta(x-y) e^{-\frac{(q-x)^2}{2\sigma^2}} dx dq \\
&= \int_B \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(q-y)^2}{2\sigma^2}} dq.
\end{aligned}$$

We have the result that the posterior distribution for q is normal with translation parameter y and fixed variance σ^2 .

Although the results are predictable in this case without going through the complicated construction given above, this example is included for two reasons. First, it is considered useful to depict the method for a continuous distribution. Second, the close relationship between the normal family and the Poisson family becomes transparent when viewed from the point of view of group theoretic methods applied to states.

4. The binomial family

We construct the binomial family similarly as was done for the Poisson family. In this case the coherent states are built from an irreducible representation of the rotation group $SO(3)$ of real 3×3 orthogonal matrices with determinant one, instead of the Weyl-Heisenberg group. The Weyl-Heisenberg Lie algebra is three-dimensional nilpotent whereas the Lie algebra corresponding to $SO(3)$ is three-dimensional simple.

4.1. The rotation group and Lie algebra

Although there are nine matrix elements in a 3×3 real matrix, the constraints of orthogonality and unit determinant for an element g of $SO(3)$, imply that g can be identified by three real parameters. There are two general methods for indicating the parameters. One way is by using the three Euler angles. The other way is to specify an axis ν and an angle t of rotation about that axis.

The rotation group is locally isomorphic to the group $SU(2)$ of 2×2 complex unitary matrices of determinant one. An element u of $SU(2)$ is identified by two complex numbers α and β where $|\alpha|^2 + |\beta|^2 = 1$. The relationship between $SO(3)$ and $SU(2)$ is that of a unit sphere to its stereographic projection upon the complex plane as shown in Naimark (1964). Although the relationship is actually homomorphic, (one g to two u), they have the same Lie algebra and so can be used interchangeably in the context presented here. It is more intuitive to work with $SO(3)$ but it is more tractable mathematically to work with $SU(2)$. Both $SO(3)$ and $SU(2)$ are compact as topological groups (Vilenkin (1968)).

In this case, since we start with matrices, basis elements of the Lie algebra can be easily obtained by differentiating the three matrices corresponding to the subgroups of rotations about the three spatial coordinate axes (x, y, z) . Thus, for example, the subgroup of $SO(3)$, indicating rotations about the z axis is given by

$$a_3(t) = \begin{pmatrix} \cos t & -\sin t & 0 \\ \sin t & \cos t & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Differentiating each matrix element with respect to t and then setting $t = 0$, we obtain the algebra basis

element

$$e_3 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Similarly, we obtain the three basis elements, e_1, e_2, e_3 with commutation relations

$$[e_1, e_2] = e_3, [e_2, e_3] = e_1, [e_3, e_1] = e_2. \quad (4.1.1)$$

4.2. A homogeneous space for the group

The rotation group $G = SO(3)$ acts as a transformation group in three-dimensional Euclidean space. However, $SO(3)$ is not transitive on the whole space. It is transitive on spheres. We take the unit two-sphere \mathbb{S}^2 as a homogeneous space for the group. Similarly as in Section 2.6, there is not a one-to-one relationship between group elements and points on the unit sphere. A one-to-one relationship (excluding the South Pole of the sphere) is provided by the cosets G/H_{NP} of the group with respect to the isotropy subgroup H_{NP} of the North Pole $(0, 0, 1)$ of \mathbb{S}^2 . In $SO(3)$, the subgroup is the group $a_3(t)$ of rotations about the z axis. In $SU(2)$, the subgroup $U(1)$ is the set of diagonal matrices $h(t)$ with diagonal elements e^{it} and e^{-it} .

Following Perelomov(1986), we consider cosets $SO(3)/SO(2)$ or cosets $SU(2)/U(1)$.

The one-to-one relationship of cosets $SU(2)/U(1)$ with the unit sphere \mathbb{S}^2 (excluding the South Pole) is given in the following manner.

Given a point ν on the unit sphere indicated by

$$\nu = (\sin \theta \cos \gamma, \sin \theta \sin \gamma, \cos \theta), \quad (4.2.1)$$

associate the coset g_ν where,

$$g_\nu = \exp \left(\frac{i\theta}{2} (\sin \gamma M_1 - \cos \gamma M_2) \right),$$

where the matrices M_1 and M_2 are given by

$$M_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad M_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.$$

In terms of rotations, the matrix describes a rotation by angle θ about the axis indicated by direction $(\sin \gamma, -\cos \gamma, 0)$ which is perpendicular to both the North Pole and the vector ν .

We can express a general element u of $SU(2)$ by

$$u = g_\nu h, \quad \text{where } g_\nu \in \text{coset } SU(2)/U(1), \quad h \in U(1). \quad (4.2.2)$$

4.3. Irreducible representations

Now we consider a representation of the algebra with basis given by linear operators E_k , for $k = 1, 2, 3$ which operate in a Hilbert space \mathcal{H} . Since the group is compact, general theory provides the result that irreducible representations correspond to finite dimensional Hilbert spaces. In the algebra representation space, the basis elements have the same commutation relations as (4.1.1). Also, we require that the operators $J_k = iE_k$ be self-adjoint with respect to the inner product of \mathcal{H} .

Similarly as in section 2.1, introduce creation and annihilation operators

$$J_+ = J_1 + iJ_2, \quad J_- = J_1 - iJ_2.$$

It is known that any irreducible unitary representation $T(g)$ of the rotation group (or of $SU(2)$), is indexed by a non-negative integer or half-integer j . The dimension of the corresponding Hilbert space is $2j + 1$.

Choose and fix the number j indicating a definite Hilbert space \mathcal{H}_j . An orthonormal basis for \mathcal{H}_j is provided by the eigenvectors ϕ_m of the self-adjoint operator J_3 which, for fixed j , has a simple discrete and finite spectrum indexed by $m = -j, -j + 1, \dots, j - 1, j$.

As operators in \mathcal{H}_j , J_+ , J_- and J_3 have creation, annihilation, and number properties similarly as in (2.2.1):

$$\begin{aligned} J_+\phi_j = 0 \quad \text{and} \quad J_+\phi_m &= \sqrt{(j-m)(j+m+1)} \phi_{m+1}, \quad \text{for } m = -j, -j+1, \dots, j-1, \\ J_-\phi_{-j} = 0 \quad \text{and} \quad J_-\phi_m &= \sqrt{(j+m)(j-m+1)} \phi_{m-1}, \quad \text{for } m = -j+1, -j+2, \dots, j, \\ J_3\phi_m &= m\phi_m, \quad \text{for } m = -j, -j+1, \dots, j-1, j. \end{aligned} \quad (4.3.1)$$

Note that in (2.2.1) there is a minimum basis vector ϕ_0 , but no maximum indicating an infinite dimensional Hilbert space. Here we have both a minimum and a maximum basis vector.

We relate ϕ_m and ϕ_{-j} by

$$\phi_m = \sqrt{\frac{(j-m)!}{(j+m)!(2j)!}} (J_+)^{j+m} \phi_{-j}. \quad (4.3.2)$$

For fixed number j , and for each $u \in SU(2)$, let $u \rightarrow T^j(u)$ denote an irreducible representation of $SU(2)$ in the Hilbert space \mathcal{H}_j , where each operator T^j is unitary with respect to the inner product in \mathcal{H}_j .

From (4.2.2), we have $T^j(u) = D(\nu)T^j(h)$ for $h \in U(1)$, where $D(\nu) = T^j(g_\nu)$,

$$T^j(g_\nu) = \exp(i\theta(\sin \gamma J_1 - \cos \gamma J_2)), \quad \text{for } 0 \leq \theta < \pi.$$

It can be shown that for $h \in U(1)$, $T^j(h)$ is a scalar factor of modulus one. Thus we focus upon $D(\nu)$.

4.4. The family of coherent states and the binomial family

We choose the element ϕ_{-j} as the fixed vector in \mathcal{H}_j . Then similarly as in Section 2, the family of coherent states, (as opposed to vectors) is given by

$$w(\nu) = D(\nu)\phi_{-j} = \exp(i\theta(\sin \gamma J_1 - \cos \gamma J_2)) \phi_{-j}, \quad \text{for } 0 \leq \theta < \pi.$$

As in Perelomov(1986), we find it convenient to re-parameterize, similarly as in Section 2.1, and use one complex parameter ξ along with the creation and annihilation operators instead of the two real angle parameters with the J_1 and J_2 operators. Thus for

$$\xi = \frac{i\theta}{2}(\sin \gamma + i \cos \gamma),$$

we have

$$D(\xi) = \exp(\xi J_+ - \xi^* J_-).$$

We seek an explicit expression for the coherent states $w(\nu)$. As in Section 2.3, the method is to factor the exponential function. The Baker-Campbell Hausdorff formula is not convenient to use in this case as it

was in Section 2.3. Instead, the Gauss decomposition of the group $SL(2, \mathbb{C})$ is used (see Perelomov(1986)). We obtain,

$$D(\xi) = \exp(\zeta J_+) \exp(\eta J_3) \exp(\zeta' J_-),$$

where

$$\zeta = -\tan(\theta/2)e^{-i\gamma}, \quad \eta = -2 \ln \cos |\xi| = \ln(1 + |\xi|^2), \quad \zeta' = -\zeta^*.$$

Finally, using (4.3.1) and (4.3.2), we obtain coherent states

$$w(\zeta) = \sum_{m=-j}^j \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}} \frac{\zeta^{j+m}}{(1+|\zeta|^2)^j} \phi_m.$$

In terms of angle parameters,

$$w(\theta, \gamma) = \sum_{m=-j}^j \sqrt{\frac{(2j)!}{(j+m)!(j-m)!}} \left(-\sin \frac{\theta}{2}\right)^{j+m} \left(\cos \frac{\theta}{2}\right)^{j-m} e^{-i(j+m)\gamma} \phi_m. \quad (4.4.1)$$

Thus, noting that the possible result values are the eigenvalues of J_3 , namely, $m = -j, -j+1, \dots, j-1, j$, we have

$$\mathcal{P}\{\text{result} = \ell, \text{ when the state is } w(\theta, \gamma)\} = |(\phi_\ell, w(\theta, \gamma))|^2.$$

Using the fact that the eigenvectors ϕ_m of J_3 are orthonormal, we have

$$(\phi_\ell, w(\theta, \gamma)) = \sqrt{\frac{(2j)!}{(j+\ell)!(j-\ell)!}} \left(-\sin \frac{\theta}{2}\right)^{j+\ell} \left(\cos \frac{\theta}{2}\right)^{j-\ell} e^{-i(j+\ell)\gamma}. \quad (4.4.2)$$

Upon taking the modulus squared, we have

$$\mathcal{P}_{w(\theta, \gamma)}\{\text{result} = \ell\} = \frac{(2j)!}{(j+\ell)!(j-\ell)!} \left(\sin^2 \frac{\theta}{2}\right)^{j+\ell} \left(\cos^2 \frac{\theta}{2}\right)^{j-\ell}, \quad \text{for } \ell = -j, -j+1, \dots, j-1, j.$$

For the binomial (n, p) distribution, put $n = 2j$, renumber the possible values by $k = j + \ell$ and put $p = \sin^2 \frac{\theta}{2}$, (the arcsine transformation).

Note that, similarly as in the Poisson and normal distributions, one of the parameters drops out upon taking the modulus squared and we are left with the one real parameter θ .

4.5. An inferred distribution on the parameter space

The parameters θ and γ index the parameter space, that is, the points of the unit two-sphere, which is isomorphic to the cosets $SU(2)/U(1)$, or equivalently $SO(3)/SO(2)$, and where points are given by the (three-dimensional) vector ν as in (4.2.1). In other words, we can take the unit sphere to be the parameter space.

The coherent states, also indexed by the point ν , are complete in the Hilbert space \mathcal{H}_j . As before, we have an isometric map from \mathcal{H}_j to the Hilbert space \mathcal{H}_{CS}^j spanned by the coherent states.

Since $D(\nu)$ takes one coherent state into another coherent state, we have the action of $D(\nu)$ on \mathcal{H}_{CS}^j . The (normalized) measure invariant to the action of $D(\nu)$ is Lebesgue measure on the sphere:

$$d(\theta, \gamma) = \frac{2j+1}{4\pi} \sin \theta d\theta d\gamma.$$

Suppose that we have an observed binomial count value \tilde{k} which, with $\tilde{k} = j + \tilde{\ell}$, gives $\tilde{\ell} = \tilde{k} - j$, for possible values $\tilde{\ell} = -j, -j + 1, \dots, j - 1, j$ corresponding to possible values $\tilde{k} = 0, 1, 2, \dots, 2j$. Then the corresponding posterior distribution on the parameter space, derived from a POV measure is

$$\mathcal{P}\{(\theta, \gamma) \in \Delta \text{ when the state is } \phi_{\tilde{\ell}}\} = \int_{\Delta} |(\phi_{\tilde{\ell}}, w(\nu))|^2 d(\theta, \gamma),$$

where the inner product inside the integral sign that of \mathcal{H}_j . From the expression for $w(\nu)$ in (4.4.1) and the fact that the vectors ϕ_m are orthonormal in \mathcal{H}_j giving the inner product (4.4.2), we have the joint distribution of θ and γ :

$$\begin{aligned} & \frac{2j+1}{4\pi} \iint_{\Delta} \left| \sqrt{\frac{(2j)!}{(j+\tilde{\ell})!(j-\tilde{\ell})!}} \left(-\sin \frac{\theta}{2}\right)^{j+\tilde{\ell}} \left(\cos \frac{\theta}{2}\right)^{j-\tilde{\ell}} e^{-i(j+\tilde{\ell})\gamma} \right|^2 \sin \theta d\theta d\gamma, \\ &= \frac{2j+1}{4\pi} \iint_{\Delta} \frac{(2j)!}{(j+\tilde{\ell})!(j-\tilde{\ell})!} \left(\sin^2 \frac{\theta}{2}\right)^{j+\tilde{\ell}} \left(\cos^2 \frac{\theta}{2}\right)^{j-\tilde{\ell}} \sin \theta d\theta d\gamma. \end{aligned}$$

For the marginal distribution of θ , we integrate γ from $0 \leq \gamma < 2\pi$ obtaining,

$$\begin{aligned} & \mathcal{P}\{\theta \in B \text{ when the state is } \phi_{\tilde{\ell}}\} \\ &= \frac{2j+1}{2} \int_B \frac{(2j)!}{(j+\tilde{\ell})!(j-\tilde{\ell})!} \left(\sin^2 \frac{\theta}{2}\right)^{j+\tilde{\ell}} \left(\cos^2 \frac{\theta}{2}\right)^{j-\tilde{\ell}} \sin \theta d\theta, \\ &= \frac{n+1}{2} \int_B \frac{n!}{\tilde{k}!(n-\tilde{k})!} \left(\sin^2 \frac{\theta}{2}\right)^{\tilde{k}} \left(\cos^2 \frac{\theta}{2}\right)^{n-\tilde{k}} \sin \theta d\theta. \end{aligned}$$

5. Discussion

We have considered a group theoretic context for three one-parameter probability families making the distinction between a *predictive* situation and a *retrodictive* one. In the predictive case, one is reasoning from a probability model for an experiment in regard to a possible result before the experiment is performed. In the retrodictive case, one is reasoning from an observed result of an experiment to a choice of a particular probability distribution within a given family that would have provided a suitable model for the experiment. Another way to distinguish between the two situations is *deductive* (reasoning from the general to the particular) as opposed to *inductive* or *inferential* (reasoning from the particular to the general). The aim is to put the predictive model in a context which enables the construction of a suitable duality relationship leading to an inferential probability distribution on the parameter space. In the Bayes context this amounts to the construction of a “noninformative” prior distribution.

It has long been supposed that group theoretic methods would lead the way to perform that kind of inference. In this paper, we have placed group theoretic methods within a context commonly used in the field of quantum mechanics. In order to do that it was necessary to construct probability distributions by the methods used in that field.

We constructed three probability families occurring commonly in the field of statistics as examples. We started from given Lie algebras relating to the field of “coherent states” and extracted the mathematical content from the quantum mechanics constructs thereby leaving the specifics of the physics behind.

Generating probability families from single fixed probability distributions as is done in quantum mechanics by using coherent states is straightforward and leads to probability distributions with the usual frequency interpretation. In fact, in quantum mechanics, many experiments are performed with “beams” of “particles” so that their results are in the form of frequency distributions.

However, setting up a dual inferential probability distribution on the parameter space is not straightforward and leads to complications and contradictions even within the field of quantum mechanics. The three examples we chose come from two Lie algebras which are of the simplest type, the nilpotent Weyl-Heisenberg algebra in the case of the Poisson and normal distributions and the “simple” rotation Lie algebra in the case of the binomial distribution. It is well known in the field of quantum mechanics that, in general, there is no frequency interpretation for the inferential probability distributions derived from POV measures. However, there is a theorem by Naimark which enables a frequency interpretation at the cost of enlarging the Hilbert space by affixing an additional random device.

From the point of view of the field of statistics we point out some features of our constructs. In regard to the relationship between the Poisson distribution and the normal translation family, note that in a sense they are two sides of the same coin in that they devolve from the same Lie algebra and the same coherent states. They are related by way of the isomorphic relationship between the \mathcal{L}^2 space of square integrable functions and the ℓ^2 space of square integrable sequences. Their parameters differ in the sense that in the normal case, the translation parameter is the α_1 parameter in the expression $\alpha = \alpha_1 + i\alpha_2$ and in the Poisson case, the parameter λ comes from the expression $\alpha = re^{i\theta}$ with $\lambda = r^2$. That is, the comparison is the one between Cartesian and polar coordinates for a point in the plane.

The relationship between the binomial and the Poisson and normal distributions is also present in terms of the coherent states. That is, the coherent states related to the Lie algebra of $SU(2)$ or $SO(3)$ approach the coherent states related to the Weyl-Heisenberg Lie algebra as the number j (that is binomial $n/2$) goes to infinity. The Lie algebras themselves have that same relationship.

In regard to the inferential probability distributions on the parameter spaces, we note that, in the binomial case, the distribution is the same as that for the variance stabilizing transformation but not for the parameter in the Poisson case.

As to the role of group representation theory, we note that the inferential distributions are marginal distributions in all three cases. The group theory applies to the full case of the joint distribution of two parameters; that is, points in the plane under Lebesgue measure in the case of the Poisson and normal distributions and points on the unit sphere under Lebesgue measure in the case of the binomial distribution. The noninformative prior measure relates to those joint distributions. Once one drops down to the marginal distribution, the group theoretic interpretation is obscured.

References

- [1] Ali, S., Antoine, J., Gazeau, J., 2000. Coherent States, Wavelets and Their Generalizations. Springer-Verlag, Berlin.
- [2] Barndorff-Nielsen, O., Blaesild, P., Jensen, J., Jorgensen, E., 1982. Exponential transformation models. Proc. Royal Soc. Lond. A.379, 41-65.
- [3] Beltrametti, E. G., Cassinelli, G., 1981. The Logic of Quantum Mechanics. Addison-Wesley, Reading, Massachusetts.
- [4] Busch, P., Grabowski, M. and Lahti, P., 1995. Operational Quantum Physics. Springer-Verlag, Berlin.
- [5] Diaconis, P., 1988. Group Representations in Probability and Statistics. Institute of Mathematical Statistics Lecture Notes, vol. 11.
- [6] Eaton, M. L., 1989. Group Invariance Applications In Statistics. Regional Conference Series in Probability and Statistics, Volume 1. Inst. Math. Stat. Hayward, California.

- [7] Good, I. J. and Gaskins, R. A., 1971. Nonparametric roughness penalties for probability densities. *Biometrika* 58, (2), 255-277.
- [8] Helland, I., 1999. Quantum mechanics from symmetry and statistical modeling. *International Journal of Theoretical Physics*, 38, 1851-1888.
- [9] Helland, I., 2003a. Quantum theory as a statistical theory under symmetry and complementarity. Submitted.
- [10] Helland, I., 2003b. Extended statistical modeling under symmetry: The link towards quantum mechanics. Submitted.
- [11] Helstrom, C., 1976. Quantum Detection and Estimation Theory. Academic Press, New York.
- [12] Kass, R. and Wasserman, L., 1996. The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, Vol 91, No. 435, 1343-1370.
- [13] Malley, J. and Hornstein, J., 1993. Quantum statistical inference. *Statistical Science*, Vol 8, No.4, 433-457.
- [14] McMurry, M.C., 1994. Quantum Mechanics. Addison-Wesley, Reading, Massachusetts.
- [15] Miller, Jr., W., 1972. Symmetry Groups and Their Applications. Academic Press, New York.
- [16] Naimark, M. A., 1964. Linear Representations of the Lorentz Group. Pergamon Press, Berlin.
- [17] Parthasarathy, K. R., 1992. An Introduction to Quantum Stochastic Calculus. Birkhäuser-Verlag, Berlin.
- [18] Perelomov, A., 1986. Generalized Coherent States and Their Applications. Springer-Verlag, Berlin.
- [19] Peres, A., 1995. Quantum Theory: concepts and Methods, Kluwer Academic Publishers, Dordrecht.
- [20] Sattinger, D. H., Weaver, O. L., 1986. Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics. Springer-Verlag. New York.
- [21] Vilenkin, N. J., 1983. Special Functions and the Theory of Group Representations. American Mathematical Society. Providence, Rhode Island.
- [22] Whittle, P., 1992. Probability via Expectation. Springer-Verlag. Berlin.