Probability on Compact Lie Groups

by David Applebaum, University of Sheffield [1]

Prepared for forthcoming INTERNATIONAL ENCYCLOPEDIA OF STATISTICAL SCIENCES, to be published by Springer

1. Introduction. Probability on groups enables us to study the interaction between chance and symmetry. In this article I'll focus on the case where symmetry is generated by continuous groups, specifically compact Lie groups. This class contains many examples such as the n-torus, special orthogonal groups $SO(n)$ and special unitary groups $SU(n)$ which are important in physics and engineering applications. It is also a very good context to demonstrate the key role played by non-commutative harmonic analysis via group representations. The classic treatise [4] by Heyer gives a systematic mathematical introduction to this topic while Diaconis [2] presents a wealth of concrete examples in both probability and statistics.

For motivation, let ρ be a probability measure on the real line. Its characteristic function $\widehat{\rho}$ is the Fourier transform $\widehat{\rho}(u) = \int_{\mathbb{R}} e^{iux}\rho(dx)$ and $\widehat{\rho}$ uniquely determines ρ. Note that the mappings $x \to e^{iux}$ are the irreducible unitary representations of $\mathbb{R}$.

Now let G be a compact Lie group and ρ be a probability measure defined on G. The group law of G will be written multiplicatively. If we are given a probability space $(\Omega, \mathcal{F}, P)$ then ρ might be the law of a G-valued random variable defined on Ω. The *convolution* of two such measures ρ_1 and ρ_2 is the unique probability measure $\rho_1 * \rho_2$ on G such that

$$\int_G f(\sigma)(\rho_1 * \rho_2)(d\sigma) = \int_G \int_G f(\sigma\tau)\rho_1(d\sigma)\rho_2(d\tau),$$

for all continuous functions f defined on G. If X_1 and X_2 are independent G-valued random variables with laws ρ_1 and ρ_2 (respectively), then $\rho_1 * \rho_2$ is the law of X_1X_2.

2. Characteristic Functions. Let $\widehat{G}$ be the set of all irreducible unitary representations of G. Since G is compact, $\widehat{G}$ is countable. For each $\pi \in \widehat{G}, \sigma \in G, \pi(\sigma)$ is a unitary (square) matrix acting on a finite dimensional complex inner product space V_π having dimension d_π. Every group has the trivial representation δ acting on $\mathbb{C}$ by $\delta(\sigma) = 1$ for all $\sigma \in G$. The *characteristic function* of the probability measure ρ is the matrix-valued function $\widehat{\rho}$ on $\widehat{G}$ defined uniquely by

$$\langle u, \widehat{\rho}(\pi)v\rangle = \int_G \langle u, \pi(\tau)v\rangle\rho(d\tau),$$

[1] Probability and Statistics Department, Hicks Building, Hounsfield Road, Sheffield, England, S3 7RH. D.Applebaum@sheffield.ac.uk

for all $u, v \in V_\pi$. $\widehat{\rho}$ has a number of desirable properties ([11]). It determines ρ uniquely and for all $\pi \in \widehat{G}$:

$$\widehat{\rho_1 * \rho_2}(\pi) = \widehat{\rho_1}(\pi)\widehat{\rho_2}(\pi).$$

In particular $\widehat{\delta} = 1$.

Lo and Ng [8] considered a family of matrices $(C_\pi, \pi \in \widehat{G})$ and asked when there is a probability measure ρ on G such that $C_\pi = \widehat{\rho}(\pi)$. They found a necessary and sufficient condition to be that $C_\delta = 1$ and that the following non-negativity condition holds: for all families of matrices $(B_\pi, \pi \in \widehat{G})$ where B_π acts on V_π and for which $\sum_{\pi \in S_\pi} d_\pi \mathrm{tr}(\pi(\sigma)B_\pi) \geq 0$ for all $\sigma \in G$ and all finite subsets S_π of V_π we must have $\sum_{\pi \in S_\pi} d_\pi \mathrm{tr}(\pi(\sigma)C_\pi B_\pi) \geq 0$.

3. Densities. Every compact group has a bi-invariant finite Haar measure which plays the role of Lebesgue measure on $\mathbb{R}^d$ and which is unique up to multiplication by a positive real number. It is convenient to normalise this measure (so it has total mass 1) and denote it by $d\tau$ inside integrals of functions of τ. We say that a probability measure ρ has a *density* f if $\rho(A) = \int_A f(\tau)d\tau$ for all Borel sets A in G. To investigate existence of densities we need the *Peter-Weyl theorem* that the set of functions $\{d_\pi^{\frac{1}{2}}\pi_{ij}; 1 \leq i, j \leq d_\pi, \pi \in \widehat{G}\}$ are a complete orthonormal basis for $L^2(G, \mathbb{C})$. So any $f \in L^2(G, \mathbb{C})$ can be written

$$f(\sigma) = \sum_{\pi \in \widehat{G}} d_\pi \mathrm{tr}(\pi(\sigma)\widehat{f}(\pi)), \tag{0.1}$$

where $\widehat{f}(\pi) = \int_G f(\tau)\pi(\tau^{-1})d\tau$ is the Fourier transform. In [1] it was shown that ρ has a square-integrable density f (which then has an expansion as in (0.1)) if and only if $\sum_{\pi \in \widehat{G}} d_\pi \mathrm{tr}(\widehat{\rho}(\pi)\widehat{\rho}(\pi)^*) < \infty$ where * is the usual matrix adjoint. A sufficient condition for ρ to have a continuous density is that $\sum_{\pi \in \widehat{G}} d_\pi^{\frac{3}{2}} |\mathrm{tr}(\widehat{\rho}(\pi)\widehat{\rho}(\pi)^*)|^{\frac{1}{2}} < \infty$ in which case the series on the right hand side of (0.1) converges absolutely and uniformly (see Proposition 6.6.1 on pp.117-8 of [3]).

4. Conjugate Invariant Probabilities. Many interesting examples of probability measures are conjugate invariant, i.e. $\rho(\sigma A \sigma^{-1}) = \rho(A)$ for all $\sigma \in G$. In this case there exists $c_\pi \in \mathbb{C}$ such that $\widehat{\rho}(\pi) = c_\pi I_\pi$ where I_π is the identity matrix in V_π [10]. If a density exists it takes the form $f(\sigma) = \sum_{\pi \in \widehat{G}} d_\pi \overline{c_\pi} \chi_\pi(\sigma)$, where $\chi_\pi(\sigma) = \mathrm{tr}(\pi(\sigma))$ is the group character.

Example 1. Gauss Measure. Here $c_\pi = e^{\sigma^2 \kappa_\pi}$ where $\kappa_\pi \leq 0$ is the eigenvalue of the group Laplacian corresponding to the Casimir operator $\kappa_\pi I_\pi$ on V_π and $\sigma > 0$. For example if $G = SU(2)$ then it can be parametrized by the

Euler angles ψ, ϕ and θ, $\widehat{G} = \mathbb{Z}_+, \kappa_m = -m(m+2)$ and we have a continuous density depending only on $0 \leq \theta \leq \frac{\pi}{2}$:

$$f(\theta) = \sum_{m=0}^{\infty} (m+1) e^{-\sigma^2 m(m+2)} \frac{\sin((m+1)\theta)}{\sin(\theta)}.$$

Example 2. Laplace Distribution. This is a generalisation of the double exponential distribution on $\mathbb{R}$ (with equal parameters). In this case $c_\pi = (1 - \beta^2 \kappa_\pi)^{-1}$ where $\beta > 0$ and κ_π is as above.

5. Infinite Divisibility. A probability measure ρ on G is *infinitely divisible* if for each $n \in \mathbb{N}$ there exists a probability measure $\rho^{\frac{1}{n}}$ on G such that the nth convolution power $(\rho^{\frac{1}{n}})^{*n} = \rho$. Equivalently $\widehat{\rho}(\pi) = \widehat{\rho^{\frac{1}{n}}}(\pi)^n$ for all $\pi \in \widehat{G}$. If G is connected as well as compact any such ρ can be realised as μ_1 in a weakly continuous convolution semigroup of probability measures $(\mu_t, t \geq 0)$. For a general Lie group, such an *embedding* may not be possible and the investigation of this question has generated much research over more than thirty years [9]. The structure of convolution semigroups has been intensely analysed. These give the laws of group-valued *Lévy processes*, i.e. processes with stationary and independent increments [7]. In particular there is a Lévy-Khintchine type formula (originally due to G.A.Hunt) which classifies these in terms of the structure of the infinitesimal generator of the associated Markov semigroup that acts on the space of continuous functions. One of the most important examples is Brownian motion and this has a Gaussian distribution. Another important example is the *compound Poisson process*

$$Y(t) = X_1 X_2 \cdots X_{N(t)} \tag{0.2}$$

where $(X_n, n \in \mathbb{N})$ is a sequence of i.i.d. random variables having common law ν (say) and $(N(t), t \geq 0)$ is an independent Poisson process of intensity $\lambda > 0$. In this case $\mu_t = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} \nu^{*n}$. Note that μ_t does not have a density and it is conjugate invariant if ν is.

6. Applications. There is intense interest among statisticians and engineers in the *deconvolution problem* on groups. The problem is to estimate the signal density f_X from the observed density f_Y when the former is corrupted by independent noise having density f_ϵ, so the model is $Y = X\epsilon$ and the inverse problem is to untangle $f_Y = f_X * f_\epsilon$. Inverting the characteristic function enables the construction of non-parametric estimators for f_X and optimal rates of convergence are known for these when f_ϵ has certain smoothness properties ([5],[6]). In [10] the authors consider the problem of *decompounding*, i.e. to obtain non-parametric estimates of the density of X_1 in (0.2) based on i.i.d. observations of a noisy version of Y: $Z(t) = \epsilon Y(t)$, where ϵ is independent of $Y(t)$. This is applied to multiple scattering of waves from complex media by working with the group $SO(3)$ which acts as rotations on the sphere.

References

[1] D.Applebaum, Probability measures on compact groups which have square-integrable densities, *Bull. Lond. Math. Sci.* **40** 1038-44 (2008)

[2] P.Diaconis, *Group Representations in Probability and Statistics*, Lecture Notes-Monograph Series Volume 11, Institute of Mathematical Statistics, Hayward, California (1988)

[3] J.Faraut, *Analysis on Lie Groups*, Cambridge University Press (2008)

[4] H.Heyer, *Probability Measures on Locally Compact Groups*, Springer-Verlag, Berlin-Heidelberg (1977)

[5] P.T.Kim, D.S.Richards, Deconvolution density estimators on compact Lie groups, *Contemp. Math.* **287**, 155-71 (2001)

[6] J-Y Koo, P.T.Kim, Asymptotic minimax bounds for stochastic deconvolution over groups, *IEEE Trans. Inf. Theory* **54**, 289-98 (2008)

[7] M.Liao, *Lévy Processes in Lie Groups*, Cambridge University Press (2004).

[8] J.T-H.Lo, S-K.Ng, Characterizing Fourier series representations of probability distributions on compact Lie groups, *Siam J. Appl. Math.* **48** 222-8 (1988)

[9] M.McCrudden, An introduction to the embedding problem for probabilities on locally compact groups, in *Positivity in Lie Theory: Open Problems* ed. Hilgert/Lawson/Neeb/Vinberg, Walter de Gruyter Berlin Ney York, 147-64 (1998)

[10] S.Said, C.Lageman, N.LeBihan, J.H.Manton, Decompounding on compact Lie groups, to appear in *IEEE Trans. Inf. Theory* - preprint available from arXiv:0907.2601.v1[cs.IT].

[11] E.Siebert, Fourier analysis and limit theorems for convolution semigroups on a locally compact group, *Advances in Math.* **39**, 111-54 (1981)