Friedman. And minimality echoes the spirit of causal models of explanation, where such models are used to rule out derivations involving irrelevant factors (such as the hexed salt example) that are deemed explanatory by the deductive-nomological model.

As I mentioned at the outset, my aim has not been to attack Lange's general thesis that proofs by mathematical induction are not explanatory, but rather to undercut the argument that Lange provides for this thesis. To defend the contrary claim that certain inductive proofs are explanatory would require providing a worked-out theory of mathematical explanation, something that I have not tried to do here. What I have tried to do is to show that there are plausible ways of thinking about mathematical explanation which are in tension with key steps in Lange's argument, and to this extent his premises are more 'controversial' than he admits.[4]

Swarthmore College
Swarthmore, PA 19081, USA
abaker1@swarthmore.edu

References

Baker, A. 2009. Mathematical accidents and the end of explanation. In *New Waves in the Philosophy of Mathematics*, eds O. Bueno and Ø. Linnebo, 137–59. Palgrave Macmillan.

Hafner, J. and P. Mancosu. 2005. The varieties of mathematical explanations. In *Visualization, Explanation and Reasoning Styles in Mathematics*, eds P. Mancosu, K.F. Jørgensen, and S.A. Pedersen, 215–50. Dordrecht: Springer.

Lange, M. 2009. Why proofs by mathematical induction are generally not explanatory. *Analysis* 69: 203–11.

Lange, M. forthcoming. What are mathematical coincidences (and why does it matter)? *Mind*.

Bangu's random thoughts on Bertrand's paradox
Darrell P. Rowbottom and Nicholas Shackel

Bangu (2010) claims that Bertrand's paradox rests on a hitherto unrecognized assumption, which assumption is sufficiently dubious to throw the

burden of proof back onto 'objectors to [the principle of indifference]' (2010: 31). We show that Bangu's objection to the assumption is ill-founded and that the assumption is provably true.

Bangu discerns a 'common structure' in some of Bertrand's paradoxes:

> One begins with a variable x ... and then one considers a scaling transformation θ such that $x' = \theta(x)$. (Bangu 2010: 31)

Speaking in these terms on Bertrand's behalf, we construe his objection to the principle of indifference thus:

(1) At least when θ is a nice function,[1] the probability of choosing at random in the interval $[c, b] \subseteq [a, b]$ should be the same as when choosing at random in the interval $[\theta(c), \theta(b)] \subseteq [\theta(a), \theta(b)]$.
(2) Yet applying the principle of indifference results in uniform probability distributions on $[a, b]$ and $[\theta(a), \theta(b)]$, when the probability that $x \in [c, b]$ is not necessarily equal to the probability that $x' \in [\theta(c), \theta(b)]$.
(3) Contradiction; therefore reject the principle of indifference.

The second premiss might be rejected: we might have an empirical situation in which the transformed variable is a causally dependent variable, in which case there is no reason to suppose the probability distribution of the dependent variable is uniform. However, the principle of indifference is supposed to be an unrestricted principle, so Bertrand's argument cannot be generally defeated by this point.

Bangu rejects the first premiss on the ground that it rests on a false, or at least undefended, assumption:

> R: If the argument of a scaling function is random in an interval, then the scaled value is random as well (in the scaled interval). (2010: 33)[2]

Bangu takes unpredictability 'to be the standard conception of randomness' and argues against R on the basis of the following scenario:

> Suppose a machine, a random number generator, picks a value x in the interval $[a, b]$. Randomness here is understood in the predictive sense: there is no way to predict what value this choice will return. The machine records it, but does not communicate it to us. We now ask whether the transformed value is also random in the interval $[a', b']$ $[=[\theta(a), \theta(b)]]$, in the same predictive sense. One might say that this is not so, as there is a crucial difference between the value of x and the

---

1 By a nice function, we mean a deterministic continuous bijection that maps intervals neatly: $\theta([x, y]) = [\theta(x), \theta(y)]$. This rules out θ being either a random function or a chaotic function.

2 The analysis of Bertand's chord length paradox in Shackel 2007, see especially pages 159–61, shows that Bangu's common structure does *not* underlie that paradox and hence that Bangu's criticism of this assumption is irrelevant to that paradox.

transformed value. While we cannot predict the value of x, we can predict the value of the transformed: we find what value has been recorded, and we scale it. So, is the transformed value random in $[a', b']$ after all? Or, is the sense of 'prediction' not sufficiently well defined to be useful here? (2010: 33)

Bangu's argument here rests on a premiss that is patently false. In a perfectly straightforward sense, namely prior to the machine selecting, the transformed value is no more predictable than the original. What he calls the transformed value being predictable is the situation of being given the value the machine picked and using that to predict, say, the square by squaring it. But, of course, in that sense of 'predictable' the value the machine picked is precisely as predictable! Take the value the machine picked as the prediction of the value the machine picked.

Allow us to spell this out more slowly. Accept that randomness should be understood in terms of unpredictability. When we consider the truth of R, what is at issue is whether if x is not predictable, $\theta(x)$ is not predictable. Bangu's argument seems directed, instead, at whether if x is not predictable but we know x,[3] $\theta(x)$ is not predictable. Now clearly, if θ is what we called a nice function, the unpredictability of x will not carry over to $\theta(x)$ once we know x. But that is irrelevant.

The truth of R can be illustrated by an elementary example. Let X be the variable that takes its values from the roll of a fair die. Let $\Phi(X)$ be the variable that is zero in the event that X is not 1, and one in the event that X is 1.[4] Φ is known. Even though the image of Φ is smaller than its domain, $\Phi(X)$ is no more predictable than X. In order to see this, consider a sequence of values of X from such a die roll and the corresponding values of $\Phi(X)$:

S_X $\quad$ 1,2,1,1,5,3,3,1,4,1,6,4,2, ... x_n ...
S_Φ $\quad$ 1,0,1,1,0,0,0,1,0,1,0,0,0, ... $\Phi(x_n)$ ...

If either of these sequences is predictable then knowing some number of the previous members of the sequence suffices for knowing the next one. Suppose that S_X is unpredictable. That means that for all n, knowing $\langle x_1, \ldots, x_n \rangle$ does not suffice for knowing x_{n+1}. Now we consider the predictability of S_Φ. We can suppose that we know $\langle x_1, \ldots, x_n \rangle$ and we know $\langle \Phi(x_1), \ldots, \Phi(x_n) \rangle$, and

3 $\quad$ This is, presumably, the import of 'we find what value has been recorded' in the above quotation from Bangu.

4 $\quad$ Note that except for the specified dependence on a die roll this could be a finite example of Bertrand's paradox. Since $X=1$ iff $\Phi(X)=1$, absent an empirical base the principle of indifference gives $P(X=1)=1/6$ and $P(\Phi(X)=1)=1/2$. We believe that Keynes's way around this kind of example, namely appeal to indivisibility, has been too easily accepted. However, this is not the place to discuss this issue further; see Shackel and Rowbottom, Manuscript. The point here is just to have a finite case that is fully analogous to standard paradoxical cases.

because S_X is unpredictable we do not know x_{n+1}. So for all we know x_{n+1} could be 1 or 6 and so $\Phi(x_{n+1})$ could be 1 or 0 and hence we do not know $\Phi(x_{n+1})$. So knowing $\langle\Phi(x_1),\ldots,\Phi(x_n)\rangle$ does not suffice for knowing $\Phi(x_{n+1})$ and S_Φ is unpredictable. Hence if X is unpredictable, $\Phi(X)$ is unpredictable, whence if X is random, $\Phi(X)$ is random.

In short, knowledge of how to map S_X onto S_Φ does not render S_Φ any more predictable than S_X, and that is all that is required for the truth of R in this example. This argument is in fact general. It neither depends on the discreteness of the event spaces nor on Φ. It applies to a wide range of functions. Its conclusion is equivalent to the assumption that Bangu rejects.[5]

Now one might instead consider 'special cases' where θ is a constant function on the random variable X.[6] Then it is true that $\theta(X)$ is predictable (given the knowledge that θ is a constant function). However, this kind of special case does not matter for Bertrand's paradox. Bertrand needs only one non-constant function to result in inconsistent probabilities: to refute him requires *no* such functions to exist. We suggest this is precisely why Bertrand and commentators on his paradox – among them, exceptional philosophers such as Keynes (1921/1963) and van Fraassen (1989) – have not mentioned the assumption that Bangu raises. The assumption is obviously true and for that reason not mentioned.

University of Oxford
10 Merton Street
Oxford OX1 4JJ, UK
darrell.rowbottom@philosophy.ox.ac.uk

University of Cardiff
Cardiff CF10 3EU, UK
shackeln@cf.ac.uk

References

Bangu, S. 2010. On Bertrand's paradox. *Analysis* 70: 30–5.

Keynes, J.M. 1921/1963. *A Treatise on Probability*. London: Macmillan.

Shackel, N. 2007. Bertrand's paradox and the principle of indifference. *Philosophy of Science* 74: 150–75.

Shackel, N. and D.P. Rowbottom. Manuscript. Objective Bayesianism and Bertrand's paradox.

van Fraassen, B.C. 1989. *Laws and Symmetry*. Oxford: Clarendon Press.

5   The conditions on the scaling functions need to be carefully stated. Niceness as defined in n.1 is sufficient but not necessary.

6   For example, in the die rolling case, the function $\Psi(x)=0$ for $1\leq x \leq 6$.