2015

Detecting Structure in Activity Sequences: Exploring the Hot Hand Phenomenon

Taleri Hammack

John Flach

Joseph Houpt

DETECTING STRUCTURE IN ACTIVITY SEQUENCES: EXPLORING THE HOT HAND PHENOMENON

Taleri Hammack
Wright State Joint Cognitive Systems Laboratory
Dayton, Ohio
John Flach
Wright State Joint Cognitive Systems Laboratory
Dayton, Ohio
Joseph Houpt
Wright State Cognitive Modeling Group
Dayton, Ohio

Can humans discriminate whether strings of events (e.g., shooting success in basketball) were generated by a random or constrained process (e.g., hot and cold streaks)? Conventional wisdom suggests that humans are not good at this discrimination. For example, Kahneman (2011) writes that “the hot hand is entirely in the eye of the beholders, who are consistently too quick to perceive order and causality in randomness. The hot hand is a massive and widespread cognitive illusion” (p. 117). Following from Cooper, Hammack, Lemasters, and Flach (2014), a series of Monte Carlo simulations and empirical experiments examined the abilities of both humans and statistical tests (Wald-Wolfowitz Runs Test and 1/f) to detect specific constraints that are representative of plausible factors that might influence the performance of athletes (e.g., learning, non-stationary task constraints).

Cooper, Hammack, Lemasters, and Flach’s (2014) research showed that various types of constraints on binary sequences (illustrated in Figure 1) could be reliably detected by both humans and statistical tests. This study examined both statistical tests and human performance on a success dependent learning constraint that was calibrated to reflect shooting percentages representative of shooting in NBA games.

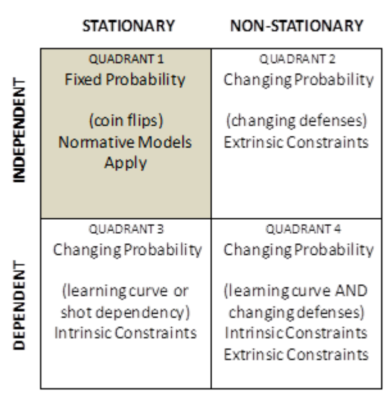

Figure 1. Types of constraints on binary sequences.

Note. This diagram illustrates four types of processes as a function of whether the generating rules are independent and stationary.

Table 1.
Monte Carlo simulation results.

	N = 1024		Runs Test			Frequency Analysis Beta Slope					
			N = 512	N = 800	N = 1024	N = 512			N = 1024		
	Mean	SD	|Z|	|Z|	|Z|	Mean	SD	*t*	Mean	SD	*t*
Quadrant 1											
Bernoulli Processes											
p(hit) = 0.3	0.306	0.013	0.04	0.18	0.44	0.02	0.06	0.87	0.02	0.04	1.50
p(hit) = 0.5	0.500	0.018	0.26	0.06	0.02	0.03	0.05	1.94	0.01	0.06	0.26
p(hit) = 0.8	0.806	0.009	0.62	0.57	0.05	-0.05	0.06	-2.58 *	-0.01	0.07	-0.63
Quadrant 2											
p(hit) = 0.2 and p(hit) = 0.6											
10% chance of alternation	0.419	0.018	3.04 *	3.87 *	4.18 *	-0.23	0.07	-10.75 *	-0.25	0.07	-11.93 *
25% chance of alternation	0.408	0.014	1.81	2.18 *	3.08 *	-0.15	0.09	-5.15 *	-0.18	0.07	-8.54 *
50% chance of alternation	0.401	0.014	0.01	0.21	0.02	0.00	0.06	-0.11	0.01	0.05	0.62
Quadrant 3											
Shot Dependencies											
Last 1 shot dependency	0.498	0.020	8.89 *	11.10 *	12.67 *	-0.55	0.12	-14.73 *	-0.56	0.08	-23.54 *
Last 5 shot dependency	0.422	0.034	3.66 *	4.40 *	5.11 *	-0.26	0.05	-15.46 *	-0.26	0.07	-11.39 *
Simple Learning Curves											
k = 0.001	0.292	0.011	0.84	2.54 *	3.79 *	-0.05	0.07	-2.30 *	-0.08	0.04	-6.94 *
k = 0.003	0.556	0.015	3.10 *	4.68 *	5.70 *	-0.14	0.08	-5.71 *	-0.10	0.05	-6.72 *
k = 0.005	0.642	0.008	4.04 *	5.33 *	5.74 *	-0.14	0.05	-9.33 *	-0.08	0.04	-6.41 *
Quadrant 4											
Simple Learning Curve and p(hit) +/-10%											
k = 0.002 +/-10%	0.391	0.022	1.98 *	2.64 *	3.18 *	-0.16	0.05	-10.23 *	-0.13	0.05	-8.28 *

Note. * p < .05 for two-tailed z-test for runs; * p < .05 for one-tailed t-test for slope (beta). N = 512 and N = 800 indicate that the beginning 512 and 800 (respectively) data points from the 1024 data point sequence were used. Includes the overall mean performance (percent success), the results of the Wald-Wolfowitz Runs Tests across sample sizes, and the mean slopes (betas) and one-tailed t-test results from the spectral analysis across sample sizes.

Table 2.
Monte Carlo simulation results for the performance dependent learning constraint.

		N = 1024		Runs Test					Frequency Analysis Beta Slope			Frequency Analysis Beta Slope		
				N = B.512	N = B.800	N = M.800	N = E.800	N = 1024	N = 512			N = 1024		
		Mean	SD	|Z|	|Z|	|Z|	|Z|	|Z|	Mean	SD	*t*	Mean	SD	*t*
Quadrant 3														
Performance Dependent Learning Curves														
p(hit) = 0.33 and p(hit) = 0.61	k = 0.001	0.380	0.019	0.19	0.16	0.48	0.51	0.40	-0.01	0.08	-0.35	-0.03	0.06	-1.37
	k = 0.003	0.459	0.022	0.33	0.64	0.48	0.30	0.66	-0.03	0.07	-1.15	-0.03	0.06	-1.60
	k = 0.005	**0.499**	**0.020**	**0.33**	**0.48**	**0.24**	**0.48**	**0.75**	**-0.05**	**0.06**	**-2.37 ***	**-0.02**	**0.05**	**-1.64**
p(hit) = 0.20 and p(hit) = 0.80	k = 0.001	0.273	0.013	0.17	0.16	0.08	0.05	0.35	-0.01	0.08	-0.40	-0.02	0.07	-0.80
	k = 0.003	0.425	0.016	0.41	1.17	0.92	1.07	2.10 *	-0.03	0.08	-1.35	-0.03	0.06	-1.86
	k = 0.005	0.529	0.022	1.31	3.11 *	2.56 *	1.81	4.09 *	-0.11	0.06	-5.63 *	-0.09	0.05	-5.53 *

Note. * p < .05 for two-tailed z-test for runs; * p < .05 for one-tailed t-test for slope (beta). B.512 and B.800 indicates the beginning 512 and 800 data points, M.800 indicates the middle 800 data points, and E.800 indicates the end 800 data points from the 1024 data point sequence. Empirical Experiment used trials generated with initial p(hit) = 0.33, asymptotic p(hit) = 0.61, and k = .005. Includes the overall mean performance (percent success), the results of the Wald-Wolfowitz

Runs Tests across sample sizes, and the mean slopes (betas) and one-tailed t-test results from the spectral analysis across sample sizes.

Method for Human Judgment Task

Participants participated in 3 blocks of thirty trials each. In each trial participants were presented with sequences representing binary strings of basketball shots (successes and misses) and were asked to discriminate between two possible generators. In the experiment participants were asked to discriminate between a sequence generated by either a Bernoulli process (a 'steady' shooter with a constant p(hit) = 0.44) or an alternative process governed by a performance dependent learning constraint with a learning rate = .005, a starting p(success) = 0.33, and an asymptote at p(success) = 0.61. Eight hundred shots were available to participants on each trial, however, the number of shots that could be viewed simultaneously decreased from 16 on Block 1 to 1 on Block 3.

Results and Discussion

Table 3.
Empirical experiment results.

	Hit Rate		False Alarm Rate		Adjusted d'		Bias Value	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Window = 16	0.64	0.03	0.33	0.05	0.82	0.17	0.04	0.09
Window = 4	0.67	0.03	0.31	0.05	0.96	0.16	0.03	0.08
Window = 1	0.61	0.04	0.34	0.03	0.70	0.13	0.06	0.07

Note. Mean percentage rate of hits and false alarms, adjusted d', and bias value as a function of window size. Mean and Standard Deviation (SD) were taken across participants.

For the Monte Carlo simulations, constrained by performance dependent learning, the Wald-Wolfowitz Runts Tests and spectral analysis (1/f) showed that none of the sequences used for the empirical experiment were detected as being significantly different from what would be expected from a Bernoulli generated process (with the exception of the spectral analysis using the weakest sample size (N = 512)). Nonetheless, there was information available in the sequences to discriminate between the two generating models as indicated by a Bayes Factor comparisons between the models fit to the generated sequences. However, when the same simulation data was used in a discrimination task done by humans, the results indicated that they *were* able to discriminate the Bernoulli generated sequences from the alternatively constrained performance dependent learning sequences significantly better than chance.

Conventional wisdom classifies the hot hand as an 'illusion' and adds it to the collection of other biases (e.g., gambler's fallacy, availability, representativeness, etc.). However, there is an alternative perspective on human reasoning that has roots in early functional/pragmatic approaches to human cognition. For example, Peirce (1877/1997) offered the construct of *abduction* as an alternative to classical logic. Abduction is an approach to rationality that is grounded in the practical success of beliefs, rather than in the syntax of arguments. More recently, an ecological rationality has been advocated by Gigerenzer (e.g., Todd and Gigerenzer, 2012). From the perspective of Ecological Rationality, heuristics are considered to be analogous to Runeson's construct of *smart instrument.* That is, the use of heuristics reflects an attunement to structure (invariants) in natural ecologies. Thus, it may be an example of an abductive form of rationality that leverages constraints in the problem ecology in ways that support successful adaptations. Thus the belief in the hot hand seems like an effective adaptive strategy rather than neglect of probability theory.

References

Cooper, J., Hammack, T., Lemasters, L. & Flach, J. (2014). Detecting Structure in Activity Sequences: Exploring the Hot Hand Phenomenon. Poster presented at the *International Society for Ecological Psychology North American Meeting*, Miami University, Oxford, OH. (June 5-7).

Kahneman, D. (2011). *Thinking fast and slow.* New York: Farrar, Straus and Giroux.

Peirce, C.S. (1877/1997). The fixation of belief. In L. Menand (ed.). *Pragmatism.* (p. 7 – 25). New York: Vintage Books.

Runeson, S. (1977). On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology,* Vol. 18. 172–179.

Todd, P.M. & Gigerenzer, G. (2012). *Ecological rationality.* New York: Oxford University Press.