No. 269

Probability Concept in the Definition of Power

by

Nozomu Matsubara

July 1985

Probability Concept in the Definition of Power

Nozomu Matsubara

Institute of Socio-Economic Planning,

University of Tsukuba, Japan

and

Yale University

Probability Concept in the Definition of Power:

Section 1. Introduction

'Power', as put by Max Weber in his Wirtschaft und Gesellschaft, is the probability that one actor within a social relationship will be in a position to carry out his own will despite resistance, regardless of the basis on which this probability rests. Not only this proposition is put on the mode of probability, but it elicits the epistemological nature of power ; Power has many ingredients (Bruce Russet [15]), but power at least as influence is probabilistic and not deterministic.

Robert Dahl's definition ([5],[6]) of power as the ability of A to get B to do something that he would otherwise not do is rightly framed in the notion of probability. A has power over B (with respect to "scopes", x, y, concerned) if it holds that

$$P \ (\text{B does y} \mid \text{A does x}) > P \ (\text{B does y} \mid \text{A does not do x}),$$

and this probability increase is defined as "amount" of power. John Harsanyi [8] asserts that one should add the notion of "strength of" power, expressed in utility concept, sufficient to make such compliance probable. See also J. Hart [9].

The concept of probability is more ubiquitous in political science. Richard Brody [3], in defining the concept "deterrence", puts; nuclear weapons make more probable (but not certain - the author) the rejection of armed aggression as a potential policy alternative. The character and the performance of any threat system heavily depend on the subjectively assessed probability, termed as "credibility", of the threat being true. (See e.g. Boulding [2])

In this ubiquity Dahl's use of the concept of probability in defining power particularly contributes to its success in eliciting the salient feature of power; it explicitly employs the mode of conditional probability, and, essentially, conditioning is tantamount to the limiting of 'freedom of choice'. Karl Deutsch [7], for who probability concept is crucial in "Nerves of Government", the study of political communication and control, recapitulates Dahl's concept of power as 'the ability to produce a change in probability distribution of a class of repetitive outcomes.' Certainly conditioning induces a change in probability distribution, but what changes?

The aim of this paper is to give a little more detail, based mainly on Dahl and Deutsch'es theories of power, to the concept and the content of power from the viewpoint of mathematical theory of probability.

Section 2. Power, Freedom of Choice, and Predictability

Let an actor, B, have a set of K alternatives { b_1, b_2, ... , b_k } from which, out of his own will or not, he is to choose, and the choice be characterized as based on { $P(b_1)$, ... , $P(b_k)$ } rather than directly on { b_1, b_2, ... , b_k }, where P stands for the probability of choice; thus the choice is thought of not deterministic, but probabilistic. Claude Shannon and Warren Weaver [17], in formulating the mathematical theory of communication, utilizes 'entropy' to define the freedom of such choice as

$$H(B) = - \sum_{i=1}^{k} P(b_i) \log P(b_i),$$

The characterizing property of H(B) is as follows; if he has (or feels) full degree of freedom, $p_1 = \dots = p_K = 1/K$, then H(B) attains maximum log K. As his choice is more concentrated or, _a forteriori_ confined to specific b, say b_i, H(B) decreases toward 0.

The typical implication for political science is found in the definition of power, where K=2 and

$$b_1 = \text{compliance } (c), \quad b_2 = \text{non-compliance } (\bar{c}),$$

and

$$H(B) = - p \log p - (1-p) \log (1-p), \quad p = P(b_1).$$

If H(B) is small, that indicates he must feel determinate (out of his own will or being coerced) to comply or not to comply, and vice versa. The case p=1/2 can admit various interpretation; full freedom of choice, complete uncertainty, or simply Laplace's so called 'principle of insufficient reason', etc. Unless otherwise stated, every member of a political body is usually supposed to have freest and independent choice. (For example, Douglas Rae [14])

Randomness and order is other fundamental aspect to differentiate human organizations, let alone political systems. (John March and Herbert Simon's [12] concept of 'absorption of uncertainty' in organizational communication reflects this point.) Ultimately it is tantamount to the freedom of choice of its members. As one is well acquainted, entropy H can qualify the state of randomness and order; if H is decreasing, order establishes, and if increasing, then declines with randomness going to prevail. What change of distribution 'power' will bring about is characterized just in this content. Let, for short, c= 'B does y' and

$$p_0 = P\,(\text{ B does y } \mid \text{ A does not do x }),$$

$$p_1 = P\,(\text{ B does y } \mid \text{ A does x }).$$

Power need not generate order, although, as one expects, it usually does. There are 3 typical cases to distinguish (Equality should not be taken per se, but approximate);

Between actors A and B

(i) order establishes, if $p_1 > p_0 \geqq 1/2$,

(ii) order declines (but for A's favor), if $1/2 \geqq p_1 > p_0$,

and

(iii) order is reserved with its favor reversed for A, if

$p_1 > 1/2 > p_0$ and $p_1 = 1 - p_0$,

since, denoting

$$H_i(B) = - p_i \log p_i - (1-p_i) \log (1-p_i), \qquad i=0,1$$

we have

$$H_1(B) < H_0(B), \quad H_1(B) > H_0(B), \quad H_1(B) = H_0(B),$$

for (i), (ii), (iii), respectively.

(When the mode of non-compliance is broken down to various responses, then we can take $K \geq 3$. In fact, Kenneth Boulding [2] categorizes responses to threat as 'submission', 'defiance', 'counter-threat' and 'integrative response'.)

However, whether $p_1 > p_0$ or $p_0 > p_1$ ('negative power'), power relationship itself introduces between A and B predictability of outcomes, c and $\bar{c}$. In fact, let

$$a = \text{'A does x'}, \quad \bar{a} = \text{'A does not do x'}.$$

Then, the probabilities of B's compliance and of non-compliance, regardless of (or ignoring) a, $\bar{a}$, are

$$P(c) = P(c|a)\,P(a) + P(c|\bar{a})\,P(\bar{a}),$$

$$P(\bar{c}) = P(\bar{c}|a)\,P(a) + P(\bar{c}|\bar{a})\,P(\bar{a}).$$

Randomness associated with outcomes of c, $\bar{c}$, are

$$- P(c) \log P(c) - P(\bar{c}) \log P(\bar{c}) = H(B)$$

when ignoring a, $\bar{a}$, and

$$P(a)\ \{ -P(c|a) \log P(c|a) - P(\bar{c}|a) \log P(\bar{c}|a) \}$$
$$+ P(\bar{a})\ \{ -P(c|\bar{a}) \log P(c|\bar{a}) - P(\bar{c}|\bar{a}) \log P(\bar{c}|\bar{a}) \}$$
$$= P(a)\ H_1(B) + P(\bar{a})\ H_0(B) = H(B|A)$$

when inserting the power consideration. THe Shannon-Weaver's celebrated result is the inequality

$$H(B) \geq H(B|A)$$

It reads 'conditioning reduces randomness and establishes order.' It can also read, in our context, 'power consideration serves predictability of outcomes.' The reduction of randomness is thus to be computed by

$$O(A:B) = H(B) - H(B|A),$$

which is the measure of order established between A and B by inserting power relationship.

One thing is worth mentioning. Let p_{ij} be joint probabilities of events (a,c) etc., like the following;

	c	$\bar{c}$
a	p_{11}	p_{12}
$\bar{a}$	p_{21}	p_{22}

Then, we have, explicitly

$$P(a) = p_{11} + p_{12} = p_{1.} , \quad P(\bar{a}) = p_{21} + p_{22} = p_{2.} ,$$

$$P(c) = p_{11} + p_{21} = p_{.1} , \quad P(\bar{c}) = p_{12} + p_{22} = p_{.2} ,$$

$$p_1 = P(c|a) = p_{11} / p_{1.} ,$$

$$p_0 = P(c|\bar{a}) = p_{21} / p_{2.} .$$

That A has power over B, i.e. $p_1 > p_0$, is equivalent to $(1/p_0) > (1/p_1)$, or

$$p_{11}\ p_{22} - p_{12}\ p_{21} > 0 ,$$

the so called determinant of the above matrix. This fact itself is interesting, which does not concern us here. More striking result is that, when a, c and $\bar{a}$, $\bar{c}$ are interchanged (with 1, 2 also interchanged), above inequality still holds and hence B has power over A !

Dahl's definition of power thus poses the problem on the symmetricity of power relation, although it is not all attributed to his definition. Rather it is grounded on more profound causality problem, as John March [11] discusses. Patrick Suppes [10] shows general causal paradoxes using a similar 2x2 probability table.

In fact, O(A:B) is also symmetrical O(A:B) = O(B:A), since one calculates to obtain

$$O(A:B) = \sum \sum p_{ij} \log (p_{ij} / p_{i.} p_{.j}),$$

which is symmetrically defined. However, order is what they share in common, symmetricity reveals less salient paradoxes.

Max Weber's definition 'despite resistance' may suggest asymmetric, unilateral relationship as inherent, but some theory, e.g. the social exchange theory (See [1]) seems to put only relative value on this statement.

If instead we would be more pragmatic in the analysis of concept of power, at least two alternatives could be proposed. One is a statistics-oriented approach, which is situated between probabilistic formalism and realistic approach. (See the next section.)

Another approach is rather akin to so called structurism, or, as its psychological counterpart, the Gestalt theory of structure and pattern. Thomas Schelling's 'focal point' ([16]) is the concept of precipitate randomness reduction, for the conflict resolution, onto that specific point just attributed to by the situation itself. Symbolically, let S be the whole pattern of situation. Then

$$w(x|S) = \begin{cases} 1 & \text{if } x=f \\ 0 & \text{otherwise} \end{cases},$$

where w is a subjective prior probability distribution over $X \ni x$ and f is a focal point. This is a very interesting type of conditioning.

Section 3. Credibility of Threats

Credibility is one of the most pragmatic probability concept that the concept of power is related to. Suppose that A, a threatener, is challenged by B, the threatened. He has two choices; to carry out doing the harm (h) and not to do that ($\bar{h}$).

Each actor can have differing assessments of the probabilities of h, $\bar{h}$; A's own and B's subjective —often called Bayesian —assessments, which we denote

$$P_A(h) = p_A , \quad P_A(\bar{h}) = 1 - p_A ,$$

$$P_B(h) = p_B , \quad P_B(\bar{h}) = 1 - p_B .$$

p_B we call 'the credibility' of A as assessed by B. The case that $p_B = p_A$ or, in statistics terminology, B is <u>well-calibrated</u> against A with respect to the crisis consciousness gives no substantial problem, though the case can be the frame of reference, being in some sense ideal one. (In statistical language, a well calibrated weather forecaster predicts by the subjective probability that coincides in the long run with actual percentage of rainy days. See Murphy and Winkler [13]. Any threatener should strategically engineer the situation so that his assessed credibility should be well-calibrated against his genuine, interior subjective probability. In this context, it may be interesting to study whether communist leaders' political announcement and their overt actions are in good calibration.)

The case $p_B \neq p_A$, or, at least they greatly differ, offers the interesting problem of misperception. Misperception in itself is not dangerous if it is consistently obeying some long run law and noticed consciously by both actors. A long run law, if any, can be put in the form with possible error ϵ,

$$\log \frac{p_B}{1-p_B} = \beta \log \frac{p_A}{1-p_A} + \alpha + \epsilon ,$$

where α is in statistics (D.R.Cox [4]) called 'bias' and β 'spread'. In short, if $\alpha > (<) 0$, p_B tends to over- (under-) estimate p_A, and if $\beta > (<) 1$, p_B to react more (less) sensitively than real to the differential change on p_A. α, β may characterize in detail the relationship between A and B.

Well-calibration of B against A can be tested by the statistical test of hypothesis that $\alpha = 0$, $\beta = 1$. Other hypotheses worth while are; $\beta = 1$, $\alpha > (<) 0$; $\beta > (<) 1$, $\alpha = 0$ etc. The logarithmic transformation above is called 'log of odds' (in favor of h over $\bar{h}$), or, in short 'logit'. The logit transformation is frequently appealed to in statistical analysis to change the domain of analysis from (0,1) to $(-\infty , \infty)$, usually by pragmatic reasons. It is worth while to note that the logit score is the rate of gain of predictability, the negative of randomness -H(B), since

$$\frac{d}{dp} (-H(B)) = \log \frac{p}{1-p} .$$

Lastly, Bayes theorem produces p_B as posterior probability, p_A being prior, as

$$\frac{p_B}{1-p_B} = \frac{f_1(z)}{f_2(z)} \frac{p_A}{1-p_A} ,$$

where z is the data (or 'signal') as the judgmental base for p_B, and f_1, f_2 are its likelihood under h and $\bar{h}$ respectively. But the above expression is equivalent to the hypothesis that $\beta=1$. Thus the test of $\beta=1$ checks B to be (or not) Bayes assessor of the situation. Many psychological researches have been accumulated to assess a judge from the viewpoint of the speed of modification of probability assessment, as contrasted with Bayesian modification. The speed is most notably found to be rather slow, known as 'conservatism'.

REFERENCES

1. Baldwin, D.A. (1978) Power and Social Exchange. American Political Science Review, 72, pp.1229-1242.

2. Boulding, K.E. (1963) Toward a Pure Theory of Threat Systems. American Economic Review, 53, pp.424-434.

3. Brody, R. (1968) Deterrence. Encyclopedia of Social Sciences. MacMillan and Free Press.

4. Cox, D.R. (1958) Two Further Applications of a Model for Binary Regression. Biometrika, 45, pp.562-565.

5. Dahl, R.A. (1957) The Concept of Power. Behavioral Science, 2, pp.201-215.

6. Dahl, R.A. (1968) Power. Encyclopedia of Social Sciences. MacMillan and Free Press.

7. Deutsch, K.W. (1966) The Nerves of Government——Models of Political Communication and Control. The Free Press.

8. Harsanyi, J.C. (1962) Measurement of Social Power, Opportunity Cost, and the Theory of Two-person Bargaining Games. Behavioral Science, 7, pp.67-80.

9. Hart, J. (1976) Three Approaches to the Measurement of Power in International Relations. International Organization, 30, pp.289-305.

10. Hintikka, J. and Suppes, P. (1966) Aspects of Inductive Logic. Studies in Logic and the Foundation of Mathematics, North-Holland.

11. March, J.G. (1957) An Introduction to the Theory and Measurement of Influence. American Political Science Review, 49, pp.431-531.

12. March, J.G. and Simon, H. (1958) Organizations. John Wiley and Sons.

13. Murphy, A.H. and Winkler, R.L. (1977) Reliability of Subjective Probability Forecast of Precipitation and Temperature. Journal of the Royal Statistical Society, C, 26, pp.41-47.

14. Rae, D.W. (1969) Decision-Rules and Individual Values in Constitutional Choice. American Political Science Review, 63, pp.40-56.

15. Russet, B. and Starr, H. (1981) World Politics: The Menu for Choice. W.H. Freeman and Co.

16. Schelling, T.C. (1976) The Strategy of Conflict. (5th printing) Harvard University Press.

17. Shannon, C. and Weaver, W. (1949) The Mathematical Theory of Communication. University of Chicago Press.