Special Issue - 2016

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
IC-QUEST - 2016 Conference Proceedings

Laws of Information in Probabilistic System

Dr. H.R. Trivedi
Associate professor, D. B. Science College Gondia
Gondia(441614), Maharashtra,
India

Dilendra C. Katre
Department of Mathematics,
M. I. Patel College Soni, Gondia,
India

Abstract:- **Famous Physicist Newton given by three laws of motion. In this paper we state three laws of information and its some implications for modern science and technology. We also discuss how these laws of information in probabilistic systems are parallel to newton's laws of motion.**

Keywords:- Probability distribution, information constraints; directed divergence; maximum entropy and minimum cross entropy; uniform distribution.

INTRODUCTION:-

Isaac Newton (1643-1727) the famous physicist who formulated the laws of motion. Three laws of motion given by,
1) Every object in a state of uniform motion tends to remain in that state of motion unless an external force is applied to it. It is also called law of inertia.
2) The force is proportional to the rate of change of momentum.
3) For every action there is an equal and opposite reaction.
The laws of information are

1. First Law of Information:-

The first law of motion gives us the definition of force as an entity which is required to change the state of rest or uniform motion in that state of motion. The first law of information stated as, "Every probability distribution is to be taken as the uniform distribution unless we give some information which is not permutationally symmetric about the outcomes and which is not satisfied when all the probabilities are equal."

When outcomes are n, then probability of each outcome is $\frac{1}{n}$ is called uniform distribution. It is denoted as U_n,

$$U_n = (\frac{1}{n}, \frac{1}{n}, \ldots, \frac{1}{n}) \tag{1}$$

Probability constraints

$$a_s \leq f_s(p_1, p_2, \ldots, p_n) \leq b_s,$$
$$s = 1, 2, \ldots, n \tag{2}$$

These constraints given from probability distribution. Information will available in various forms
in natural constraints are

$$\sum_{i=1}^{n} p_i = 1 \qquad p_1, \ldots, p_n \geq 0 \tag{3}$$

In moment equality constraints

$$\sum_{i=1}^{n} p_i g_t(x_i) = a_t \qquad t = 1, 2, \ldots, n \tag{4}$$

In form of inequalities on probabilities $a_i \leq p_i \leq b_i$
$i = 1, 2, \ldots, n \tag{5}$

In form of a Priori probability distribution
$$Q = (q_1, q_2, \ldots, q_n) \tag{6}$$

When constrants does not change then this is called permutationally symmetric. $p_1 + p_2 + p_3 + \cdots + p_i = 1$ is symmetric & consistent.

$$p_1, \ldots, p_n \geq 0 \quad \& \quad a \leq p_i \leq b \tag{7}$$

In $k = \frac{1}{n}$ then $p_1^2 + p_2^2 + \cdots + p_n^2 = k \tag{8}$

If $a \leq f\left(\frac{1}{n}\right) \leq b$ then $a \leq f(p_1) + f(p_2) + \cdots + f(p_n) \leq b$ is consistent. $\tag{9}$

If $a \leq \frac{1}{n} \leq b$ then $a \leq p_i \leq b$ are consistent. $\tag{10}$

When the constraint may be consistent with uniform distribution but it may not be permutationally symmetric.

$$a_1 p_1 + a_2 p_2 + \cdots + a_n p_n = \frac{a_1 + a_2 + \cdots + a_n}{n} \tag{11}$$

The constraint

$$a_1 p_1 + a_2 p_2 + \cdots + a_n p_n = k \neq \frac{a_1 + a_2 + \cdots + a_n}{n} \tag{12}$$

are not consistent and not permutationally symmetric.
If an information constraint is both symmetric and consistent with uniform distribution the distribution need not be uniform. The condition of an information constraint being permutationally symmetric is neither necessary nor sufficient for the distribution to be uniform since the uniform distribution can satisfy non-symmetric constraint.
Thus, the first law of information gives definition of information constraints (a priory, equality, inequality, moment probability distribution) which are required to prevent the distribution from being uniform. There are some constraints which are 'harmless' or 'redundant', which do not prevent the distribution from begin uniform.

2. Second Law of Information:-

The second law of motion given as the force is proportional to the rate of change of momentum. All constraints which are neither symmetric nor consistent with uniform distribution are 'genuine' constraints because the force of distribution to depart from uniform distribution. The force of these constraints depends on their degrees of asymmetry and degrees of inconsistency. We can sketch our inspiration for machines, there force is not measured directly but by the rate of change of momentum it produces.

Thus the second law of information stated as, "The amount of information given by a constraint imposed on a probability distribution is measured by the excess of entropy of the uniform distribution over the maximum entropy of all distributions which are consistent with the given constraints, the constraints which are consistent with the uniform distribution contain no information."

Special Issue - 2016

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
IC-QUEST - 2016 Conference Proceedings

A measure $D(P:Q)$ for measuring directed divergence of probability distribution P from Q . $D(P:Q)=0$ iff $P=Q$ where $D(P:Q)$ is a convex function of $p_1, \ldots, p_n$ (13),

$$D(P:Q)=\sum_{i=1}^{n} p_i \log \frac{p_i}{1/n} \quad (14),$$

    A force when acting on a body of given mass produces a unique rate of change of momentum. If we have a die which has six faced, p_i –denotes probability, if the given constraint the mean number of points is **4.5**.

$$p_1 + 2p_2 + 3p_3 + 4p_4 + 5p_5 + 6p_6 = \textbf{4.5} \quad (15),$$

then we have a different directed divergence from the uniform distribution,

$$U_6 = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right) \quad (16),$$

According to the principle of minimum cross entropy of Kullback[9],

$$D(P:U) = \sum_{i=1}^{n} p_i \log \frac{p_i}{1/n} = \log n + \sum_{i=1}^{n} p_i \log p_i$$

$$= \log n - \left(-\sum_{i=1}^{n} p_i \log p_i\right) \quad (17),$$

$$S(P) = -\sum_{i=1}^{n} p_i \log p_i \quad (18),$$

    <u>Where</u> $S(P)$- is called Shannon's entropy [1],

At last $I = \min D(P:U)$ (19),

satisfying the given constraints and natural constraints as the amount of information given by the constraints.

While for Newton's second law, we needed the concepts of momentum and rate of change of momentum here we need the concept of maximum entropy of probability distribution. We also need the Jaynes [2] principle of minimum cross entropy to get a unique probability distribution to a given constraint.

3. Third Law of Information:-

    Third law of motion given as, 'For every action there is an equal and opposite reaction' then the third law of information stated as, "To every asymmetric constraint set , there corresponds a unique asymmetric probability distribution and to every asymmetric probability distribution there corresponding a unique constraint set."

Given probability distribution P and C, we can find distribution Q which minimize $D(P:Q)$. However Q will be unique up to a family of distributions.

If P,Q and C are given, we determine a measure $D(P:Q)$ which will be minimized by P for given Q and C. This measure will be unique up to monotonic increasing function.

If Q -is known to be uniform distribution then instead of the minimum cross entropy distribution.

If the constraint set is symmetrical and consistent with the uniform distribution, the minimum cross entropy probability distribution is uniform.

 Hence, every asymmetric constraint set there correspond a unique asymmetric probability distribution and to every asymmetric probability distribution there corresponding a unique constraint set.

REFERENCES:-

[1] C. E. Shannon (1948), "A mathematical theory of communication." Bell system Tech., J.,vol. 27, pp. 379-423, 623-659.

[2] E. T. Jaynes (1957), "Information theory and Statistical Mechanics." Physical reviews, vol. 106, pp 620-630.

[3] H. K. Kesvan and J. N. Kapur (1989), "The generalized Maximum Entropy Principle", IEEE.

[4] I. Csiszar, "Information theory." Academic press, New York to be published.

[5] J. Aczel and Z. Daroczy, "On measures of information and their characterizations", Academic press New York.

[6] J. N. Kapur (1990), "Maximum Entropy Models in Science and Engineering", Wiley Eastern, New Delhi and John Wiley, New York.

[7] Pl. Kannapan, "On shannon's entropy, Directed divergence and inaccuracy."

[8] R. S. Verma, "Entropy in Information Theory." Ganita 16.

[9] S. Kullback (1959), "Information Theory and Statistics", John Wiley, New York.