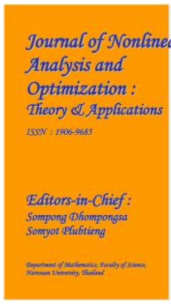
Journal of **N**onlinear **A**nalysis and **O**ptimization
Vol. 15, Issue. 1, No.3 : 2024
ISSN : **1906-9685**



JOINT AND MARGINAL DISTRIBUTION OF HYPEREDGES OVER A RANDOM HYPERGRAPH

Nabajyoti Baro Department of Mathematics, M. C. College, Barpeta, 781301, India
naba8jyoti@gmail.com
Saifur Rahman Department of Mathematics, Jamia Millia Islamia, New Delhi- 110025, India
saifur.rahman@rgu.ac.in

Abstract
Hypergraph and random hypergraph (special kind of hypergraph) are very important topic to study in current scenerio. So, paucity of studies in develop- ment of theory of random hypergraph inspire us to do a deeper study on the theory of development of hypergraph in different aspects of distribution theory. In this study, we describe the joint distribution and marginal distribution for a particular mathematical problem by forming hyperedges E on a random hypergraph H.

Keywords:
Hypergraph, Random hypergraph, Joint distribution
2020 Mathematics Subject Classification: 05C85, 05C38, 05C05, 60E05

1 Introduction

Problems related to real-life are accompanied by various unseen factors, affecting the standards of the solution and making things complex. For instance, rapid increase in the data set like, IoT (Internet of Things) like Facebook, Twitter, etc. (also called Small World [11]) work on the fundamental of a diverse user base that allows people to connect globally, recommendation system with uncertainty of liking or disliking, decision-making processes (user behaviour and trends), traffic management problem with rapid increase of automobiles, etc. One way of dealing with such unprecedential changes in the data set can be dealt by formulating them in the form of graphs (random graphs) or hypergraphs (random hypergraphs) and by leveraging the concept of graph and hypergraph, optimization can be done.

It is worth mentioning that, the problem of such complexity in graph was first realised by Paul Erdos in 1959 [13] (write briefly about what actually he studied there). Interestingly, the concept of random graph have been extensively studied and applied in various network problems in the past as well as in the recent time [15]. However, graphs (random) as a tool for modelling are extensively used in various fields and being actively studied by researchers, they only support pairwise relationships between the vertices [12]. In fact, there are many real-world problems where interactions among the objects may not be always pairwise that is, more often muti-adic interactions of the objects in group wise [2, 12].

It is to be noted that, complexity pertaining due to multi-adic relationship in a systems, may be modelled better by hypergraphs [22]. A hypergraph is a generaliza- tion of a graph introduced by Claude Jacques Berge in 1973, as a means to generalise the approaches of graphs preserving the multi-adic relationship of the objects [20]. Additionally, hypergraph theory has established itself as a brand-new, active field of study with wide range of real-world applications, like in the fields of medical science, decision-making, sociology, epidemiology, criminology, etc. Moreover, many funda- mental concepts like subhypergraphs, directed hypergraphs, weighted hypergraphs, Eulerian hypergraphs, etc. in graphs have been generalized to hypergraph theory, and their different well-known properties have been studied [18, 1, 10, 24] .

In our study, all hypergraph are consider to be simple hypergraph and occurrence of repeated hyperedges does not arise. Like graphs, hypergraphs may be classified by distinguishing between undirected and directed hypergraphs. In our study, we consider both the hypergraph, directed and undirected.

A directed hypergraph is a hypergraph with directed hyperedges. Giorgio et al.[16] discuss various properties and theorems of directed hypergraph. They de- fined the concepts of paths, cuts and connections of a hypergraph. Also they solved many problems related to operations research and computer science taking directed hypergraph as a tool. In 2018, Javidian et al.[17] also propose a model of Bayesian hypergraphs which is a form of directed acyclic hypergraph in probabilistic graphical manner. They introduce global, local and pairwise Markov properties of Bayesian hypergraphs and also define a operator named as shadow that is a projection oper- ator which maps Bayesian hypergraphs to chain graphs. An undirected hypergraph $H = (V, E)$ consists a pair of set where V is the vertices or nodes set and E is the edges or hyperedges set. Each hyperedge $e \in E$ may contain arbitrarily many vertices, the order being irrelevant, and is thus defined as a subset of V. On the other hand, hypergraph can be formed in the real life multi-criteria decision making problem. Rahman et al.[21] discussed about Choquet integral operator applied to the random hypergraph to determine the overall preference order of alternatives in practical multi-criteria decision-making problems.

The structure of this paper can be summarized as follows: Section 2 includes basic terminologies of hypergraph theory, that have been used directly or indirectly throughout the paper. In Section 3, the concept of probabilistic random hypergraphs has been discussed with a suitable examples. Next we have discussed the joint distri- bution probability and marginal distribution probability on a probabilistic random hypergraph through a particular mathematical problem on decision making. The article ends with a concluding section.

2 Preliminaries

2.1 Hypergraph

According to Berge [5], a hypergraph $H = (V, E)$ on a finite set of vertices $V = \{v_1, v_2, \cdots, v_n\}$ (called as vertex set) is defined as a family of hyperedges $E = \{e_1, e_2, \cdots, e_m\}$, where each hyperedge e_i, $i = 1, 2, \cdots, m$ is a non-empty subset of V such that $\bigcup_{i=1}^{n} e_i = V$. The order of the hypergraph is defined as the cardinality of the vertex set V and its size is defined as the cardinality of the edge set E.

Any two vertices in a hypergraph are said to be adjacent if there is a hyperedge containing both of the vertices. Whereas, two hyperedges in a hypergraph are said to be incident if their intersection is not empty. Morover, if the family of hyperedges satisfies $i \Longleftrightarrow ei \neq ej$, then H is said to be a hypergrph without repeated hyperedge. Furthermore, if $e_i \subset ej \Longrightarrow i = j$, then the hypergraph H is said to be Sperner family or simple hypergraph. In this paper, all hypergraphs are considered to be simple hyperegraph. For more general information on hypergraphs, we refer [**?**]. The degree of a hyperedge e is the number of vertices contain in e. A hypergraph is said to be k-regular if all vertices have the same degree $k \geq 0$ and H is called as r-uniform if all hyperedges have the same degree $r \geq 0$.

A hypergraph $H_1 = (V_1, E_1)$ is said to be a subhypergraph of a hypergraph $H = (V, E)$, if $V_1 \subseteq V$ and $E_1 \subseteq E$. Moreover, the subhypergraph H_1 is said to be partial subhypergraph of H if $V_1 = V$. While, the subhypergraph H_1 is said to an induced subhypergraphs of H if the hyperedges induced by V_1 in H is also contained in H_1.

Example 2.1. Let $V = \{x_1, x_2, \cdots x_{17}\}$ be a finite set of objects and $E = \{e_1, e_2, \cdots, e_7\}$, where $e_1 = \{x_1, x_2, x_3, x_{15}\}$, $e_2 = \{x_3, x_4, x_5, x_6\}$, $e_3 = \{x_6, x_7, x_8, x_9\}$, $e_4 = \{x_8, x_9, x_{10}, x_{11}, x_{13}\}$, $e_5 = \{x_{11}, x_{12}\}$, $e_6 = \{x_9, x_{13}, x_{14}, x_{15}\}$, $e_7 = x_{16}$. be the non-empty subsets of V. Then, $H = (V, E)$ is a hypergraph (see Figure 1).

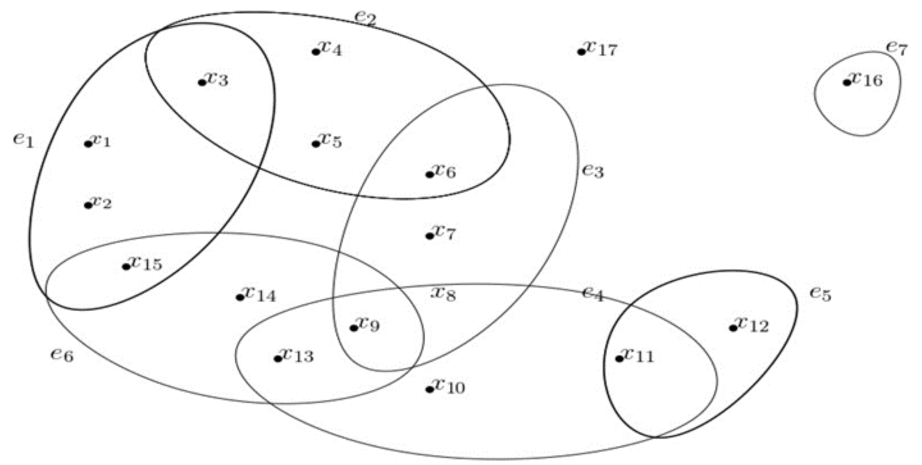


Figure 1: Hypergraph representing Example 2.1.

Observe that, the rank of the hypergraph H is $r(H) = 5$ (degree of e_4), and its co-rank is $cr(H) = 1$ (degree of e_7). It is to be noted that, the sub hypergraph $H_1 = (V_1, E_1)$, where $E_1 = \{e_1, e_2, e_3\}$) is a partial hypergraph generated by e_1, e_2, and e_3.

Now consider, the sub hypergraph $H_2 = (V_2, E_2)$ of H, such that

$$V_2 = \{x_2, x_3, x_6, x_7, x_{10}, x_{11}, x_{13}, x_{14}, x_{15}\}$$

and

$$E_2 = \{e'_1, e'_2, e'_3, e'_4\}$$

where $e'_1 = e_1 \cap V_2 = \{x_2, x_3, x_{15}\}$, $e'_2 = e_2 \cap V_2 = \{x_3, x_6\}$, $e'_3 = e_3 \cap V_2 = \{x_6, x_7\}$, and $e'_4 = e_6 \cap V_2 = \{x_{13}, x_{15}\}$) is an induced subhypergraph. It is to be noted that, the hyperedge $e' = e_5 \cap V' = \{x_{11}\}$ is not a herperedge of this induced subhypergraph.

2.2 Probabilistic random Hypergraphs

Let X be a random experiment and let S be the sample space of X. Let, $S = \{v_1, v_2, ..., v_n\}$ sample space be finite and let E be a collection of events of S, that is, $E \subseteq P(X)$, where $P(X)$ is the power set of X. Then, $G = (S, E)$ forms a hypergraph. Let the edges be $\{e_1, e_2, ..., e_n\}$. We regard that $e_1 \subseteq E$ is also an event with probability $f(e_1) = w_1$ (say), where, $0 \leq w_1 \leq 1$. Also, $f(e_i) = w_i$, for $i = 1, 2, ..., n$ and $0 \leq w_i \leq 1$. Let w_1 denote the weight of the hyperedge, then the ordered triplet (S, E, W) is a weighted hypergraph, where $W = \{w_1, w_2, ..., w_n\}$. Then, the hypergraphs (S, E), where the vertex set is the outcome of random experiment and edge set is the collection of events called random hypergraph and the weighted hypergraph (S, E, W), where weights are the corresponding probability of the edges is called probabilistic hypergraph. For details, we refer the reader to see citeRahman26.

As an illustration to the above discussion, the following example is considered:

Example 2.2. Consider the problem of assessment of the best science student from a group of 10 short-listed students $V = \{v_1, v_2, \ldots, v_{10}\}$. Suppose the assessment of the students are made on the basis of the following five criteria C_j $(j = 1, 2, 3, 4, 5)$:

1. C_1: Attendance
2. C_2: Knowledge
3. C_3: Conception
4. C_4: Computational Competence
5. C_5: Other activity

The Annual performance grades of each student v_i $(i = 1, 2, \ldots, 10)$ is given by score function values $f(v_i) = \langle v(i,1), v(i,2), v(i,3), v(i,4) \rangle$, where the partial evaluations $v(i,j)$ corresponds to the assessment grade of each student v_i $(i = 1, 2, \ldots, 10)$ with respect to each criteria C_j $(j = 1, 2, 3, 4, 5)$. Assume that the assessment grade values are measured in a scale of grade 10. Then the assumed assessment of the students in terms of the annual performance grades is expressed by the decision matrix $D = [v_{ij}]_{10\times5}$ as shown below in Table 1.

Criteria/Alternative	C_1	C_2	C_3	C_4	C_5
v_1	7.45	8.51	7.37	9.32	5.8
v_2	7.72	8.47	8.53	8.43	6.6
v_3	8.23	8.13	7.48	8.32	7.8
v_4	7.31	7.63	6.85	7.41	8.9
v_5	7.21	6.24	6.61	7.28	5.9
v_6	9.14	8.72	9.11	9.51	7.8
v_7	8.23	7.15	9.74	9.33	7.3
v_8	8.43	9.34	9.32	9.48	6.8
v_9	8.45	7.81	9.23	8.68	7
v_{10}	9.78	8.91	8.64	9.34	8.8

Table 1: The assessment decision matrix D.

1. Suppose a decision maker identify four groups e_k $(1 \le k \le l = 5)$ as the partial evaluations of the alternatives v_i $(i = 1, 2, \dots, 10)$ corresponding to the interrelated criteria from the set of criteria $C = \{C_1, C_2, C_3, C_4, C_5\}$ based on his expertise knowledge of the subject. We consider $e_1 = \{C_1, C_2\}, e_2 = \{C_2, C_4\}, e_3 = \{C_1, C_4\}, e_4 = \{C_3, C_4, C_5\}$ be four hyperedge. This will form a random hypergraph $H = (V, E)$ for the hyperedges $E = \{e_1, e_2, e_3, e_4\}$, corresponding to the criteria C_i, $i = 1, 2, \dots, 5$. The graphical representation of the random hypergraph in this case can represented in the following manner (see Figure 2).

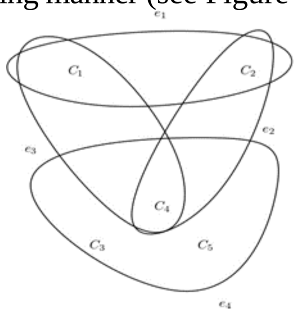


Figure 2: Random hypergraph with four hyperedges.

2. Define an interaction matrix B as follows:

$$B = [b_{ij}]_{n\times m} ,$$

where $b_{ij} = v_{ij} l_j$ $(i = 1, 2, \dots, n; j = 1, 2, \dots, m)$ and l_j $(j = 1, 2, \dots, m)$ are the interaction levels of C_j $(j = 1, 2, \dots, m)$, respectively. The interaction level of any vertex in a hypergraph refers to the number of vertices that have direct links with that vertex. For example, if a vertex v is incident with exactly t hyperedges, then the level of the vertex is determined by the number of distinct vertices contained in those t hyperedges.

Criteria/Alternative	C_1	C_2	C_3	C_4	C_5	Sum
v_1	14.9	17.02	7.37	27.96	5.8	73.03
v_2	15.44	16.94	8.53	25.29	6.6	72.8
v_3	16.64	16.26	7.48	24.96	7.8	72.96
v_4	14.62	15.26	6.85	22.23	8.9	67.86
v_5	14.42	12.48	6.61	21.84	5.9	61.25
v_6	18.28	17.44	9.11	28.53	7.8	81.16
v_7	16.46	14.3	9.74	27.99	7.3	75.79
v_8	16.86	18.68	9.32	28.44	6.8	80.1
v_9	16.9	15.62	9.23	26.04	7	74.79
v_{10}	19.56	17.82	8.64	28.02	8.8	82.84

3. We get the decision Matrix by normalization of interaction matrix.
we consider the hyperedges as $e_1 = \{C_1, C_2\}$, $e_2 = \{C_2, C_4\}$, $e_3 = \{C_1, C_4\}$, $e_4 = \{C_3, C_4, C_5\}$. we can obtained the e_i's ($i = 1, 2, \ldots, 4$) by summing up with respective C_i, ($i = 1, 2, \ldots, 5$) and dividing with the sum of respective v_i, $i = 1, 2, ..., 10$.

Criteria/edege	e_1	e_2	e_3	e_4
v_1	0.218480493	0.244079398	0.229568789	0.307871321
v_2	0.22239011	0.207005495	0.196703297	0.323626374
v_3	0.224232456	0.225466009	0.226836623	0.323464912
v_4	0.220159151	0.221632773	0.216917182	0.341290893
v_5	0.219591837	0.220734694	0.236571429	0.323102041
v_6	0.220059142	0.224618038	0.229793001	0.325529818
v_7	0.202929146	0.217442934	0.231692835	0.347935084
v_8	0.22184769	0.234956305	0.223595506	0.319600499
v_9	0.217408744	0.220484022	0.229041316	0.333065918
v_{10}	0.225615645	0.220304201	0.230806374	0.323273781

3 Joint and marginal distribution over random hyper- graph

In this section of the paper, we have studied the joint distribution on random hy- pergraph by considering two random variables X and Y, where $X = f(e_i)$ and
$Y = g(e_i)$ $e_i \in E$, ($i = 1, 2, \ldots, 4$), respectively.

$X = f(e_i)$	0.218480493	0.244079398	0.229568789	0.307871321
$Y = g(e_i)$	0.22239011	0.232142857	0.221840659	0.323626374

Now let us define the joint probability mass function of the two random variables
X and Y as

$$P_{XY}(x, y) = P(X = x, Y = y).$$

Also

$$P_{XY}(x, y) = P(X = x, Y = y) = P((X = x) \text{ and } (Y = y)).$$

We can define the joint range for X and Y as follows:

$$R_{XY} = \{(x, y)|P_{XY}(x, y) > 0\}.$$

In particular if $R_X = \{x_1, x_2, ...\}$ and $R_Y = \{y_1, y_2, ...\}$, then we can write

$$R_{XY} \subset R_X \times R_Y = \{x_i, y_j)|x_i \in R_X, y_j \in R_Y\}.$$

Moreover, for two random variables X and Y, we have

$$\sum_{(x_i,y_j)\in R_{XY}} P_{XY}(x_i, y_j) = 1$$

Thus, for any set $A \subset R^2$, where $(X, Y) \in A$, we can find the joint probability mass function as

$$\mathrm{P(X,Y)} = \sum\nolimits_{(x_i,y_j)\in R_{XY}} P_{XY}(x_i, y_j)$$

and that can be determined as follows:

p(X=x,Y=y)	Y_1	Y_2	Y_3	Y_4	∑ F(Y)
X_1	0.048587901	0.050718686	0.048467857	0.07070605	0.218480493
X_2	0.054280844	0.056661289	0.054146735	0.07899053	0.244079398
X_3	0.051053828	0.053292754	0.050927691	0.074294515	0.229568789
X_4	0.068467537	0.071470128	0.068298377	0.099635279	0.307871321
∑ F(Y)	0.22239011	0.232142857	0.221840659	0.323626374	1

3.1 Marginal distribution over random hypergraph

The joint probability mass function contains all information regarding the distributions of X and Y. We can obtain the probability mass function of X from its joint probability mass function with Y. We can write

$$P_X(x) = P(X = x) = \sum_{y_j \in R_Y} P(X = x, Y = y_j) = \sum_{y_j \in R_Y} P_{XY}(x, y_j)$$

Similarly, we can obtain the marginal probability mass function of Y as

$$P_Y(Y = y) = \sum_{x_i \in R_X} P_{XY}(x_i, y)$$

The marginal distribution of X in the Example 2.2 can be formulated by Table 2 as shown below.

X	X_1	X_2	X_3	X_4
$\sum F(X)$	0.218480493	0.244079398	0.229568789	0.307871321

Table 2: The assessment matrix Q.

Similarly, the marginal distribution of Y can be formulated as in Table 3 given below.

Y	Y_1	Y_2	Y_3	Y_4
L- $F(Y)$	0.22239011	0.232142857	0.221840659	0.323626374

Table 3: The assessment matrix Q.

4 Discussion and conclusion

We discussed about various types of hypergraph and given suitable example for degree distribution to the random hypergraph. We consider a numerical problem with ten alternatives and five criteria which forms a decision matrix. To get the interaction matrix we apply some interaction to the decision matrix. We form a random hypergraph with four hyperedges after normalising the interation matrix. Two random variables are introduced which is a function of hyperedges. These two random variables verify some distributive properities such as joint distribution and marginal distribution. This findings could be benificial for other distribution like binomial, poisson and normal distribution in future studies.

5 Ethical Statement

This article does not contain any studies with human participants performed by any of the authors.

6 Funding Statement

No fund was received for this research work.

References

[1] Bahmanian, M. A. and Sajna, M., Connection and separation in hypergraphs, *Theory and Applications of Graphs*, 2(2), (2015).
[2] Bai, S., Zhang, F., and Torr, P. H. (2021). Hypergraph convolution and hyper- graph attention. Pattern Recognition, 110, 107637.
[3] Beaumont, G.P., *Probability and Random Variables*, Ellis Horwood Limited, 1st Edition (1986).
[4] Beeri, C., Fagin,R., Maier, D. and Yannakakis, M., *On the desirability of Acyclic Database schemes*, Journal of the Association for Computing Machinary, 30, 479-513, (1983).
[5] Berge, C., *Graphs and Hypergraphs*, North Holland Publishing Co.,Amsterdam, (1973).
[6] Bhat, B. R., *Modern Probability Theory- An Introductory Text Book*, New Age International (P) Ltd., 3rd Edition, (1999).
[7] Biswas, D., *Probability and Statistics*, New Age International(P) Ltd.,(1999).
[8] Chowdhury, K. and Das, G., *Some space-biased aspects of near-rings and near- ring groups*, Int. J. Modern Math., 2(1), 103-124, 2007.
[9] Deo, N., *Graph Theory and its Applications to Engineering and Computer Sci- ence*, PHI Learning Pvt. Ltd., 1st Edition, (1974).

[10] Dewar, M., Pike, D., and Proos, J., Connectivity in hypergraphs, *Canadian Mathematical Bulletin*, 61(2), 252-271, (2018).

[11] Dodds, P., Muhammed, R., and Watts, D. J., An Experimental Study of Search in Global Social Networks, *Science*, 301, 827–829, (2003).

[12] Doley, R., Rahman, S., and Das, G., On knot separability of hypergraphs and its application towards infectious disease management, *AIMS Mathematics*, 9982- 10000, (2023).

[13] Erdos, P., Graph Theory and Probability, *Canadian Journal of Mathematics*, 11, 34-38 (1959).

[14] Feller, W., *An Introduction to Probability Theory and its Applications*, Volume I, John Willey and Sons, 3rd Edition, (2009).

[15] Frieze, A., and Karonski, M., In Introduction to Random Graphs, *Cambridge: Cambridge University Press*, (2015).

[16] Giorgio Gallo, Giustino Longo, Stefano Pallottino, Sang Nguyen, *Directed Hy- pergraphs and Applications*, Discrete Applied Mathematics, Vol. 42, 177-20, 1993.

[17] Javidian, M.A.,Lu, L, Valtora, M. and Wang, Z., *On hypergraph probabilis- tic graph model*, 2nd International Symposium on Computer and Artificial In- teligence(ISCSAI)(2018).

[18] Lonc, Z. and Naroski, P., On tours that contain all edges of a hypergraph, *The electronic journal of combinatorics*, R144-R144, (2010).

[19] Noh, H., Hickman, M. and Khani, A., *Hyperpaths in Network Based on Transit Schedules*, Transportation Research Record, Vol. 2284, 29-39, 2012.

[20] Ouvrard, X., Hypergraphs: an introduction and review, *arXiv preprint arXiv:2002.05014*, (2020).

[21] Rahman, S.,Baro, N., Kakati, P., Choquet integral operator over random hy- pergraph and its application in multicriteria decision making, preprint, DOI: 10.21203/rs.3.rs-3398945/v1, (2023)

[22] Wolf, M. M., Klinvex, A. M., and Dunlavy, D. M. (2016, September). Advan- tages to modeling relational data using hypergraphs versus graphs. In 2016 IEEE High Performance Extreme Computing Conference (HPEC) (pp. 1-7). IEEE.

[23] Zenil, H., Kiani, N. A., and Tegnér, J, A review of graph and network complex- ity from an algorithmic information perspective, *Entropy*, 20(8), 551, (2018).

[24] Zhou, D., Huang, J., and Schölkopf, B., Learning with hypergraphs: Clustering, classification, and embedding, *In Advances in neural information processing systems*, 1601–1608, (2007).