IJBPAS, June, 2023, 12(6): 2682-2691

ISSN: 2277–4998

International Journal of Biology, Pharmacy and Allied Sciences (IJBPAS)

'A Bridge Between Laboratory and Reader'

www.ijbpas.com

A REVIEW ON THE APPLICATION OF PROBABILITY DISTRIBUTION IN HEALTHCARE

PANDA NR[1, 2] PATI JK[2] AND BHUYAN R[3*]

[1] Department of Medical Research, IMS & SUM Hospital, SOA Deemed to be university Bhubaneswar, Odisha, India

[2] Department of Mathematics, CV Raman global university, Bhubaneswar, Odisha, India

[3] Department of Medical Research, IMS & SUM Hospital, SOA Deemed to be university Bhubaneswar, Odisha, India

[*]**Corresponding Author: Dr. Nihar Ranjan Panda: E Mail: niharpanda1994@gmail.com**

Received 15th July 2022; Revised 20th Aug. 2022; Accepted 5th Oct. 2022; Available online 1st June 2023

https://doi.org/10.31032/IJBPAS/2023/12.6.7191

ABSTRACT

Healthcare decision-making partially depends on the probability of disease. Present recommendations for the supervision of frequent diseases are based gradually more on scores that use arbitrary probability**.** Most of the data in the healthcare system can be described by a binomial or Poisson distribution (categorical variables). As per the previous study, normal probability distribution has a wide range of applications all most every field and especially in healthcare management. The main objective of this review is the utilization of probability distribution in the healthcare management system. Here in this paper, the author wants to illustrate how probability is helping us in a healthcare system to take necessary action based on past events and past probability distribution. Our review will put light on the healthcare system and the application of probability distribution to solve the real-life problem in healthcare analysis and healthcare management.

Keywords: Probability distribution, Binomial distribution, Poisson distribution, Healthcare

INTRODUCTION

Nowadays, in India, probability and statistics are frequently used tools in both health care and undergraduate medical courses for healthcare analysis and healthcare management like in the UK The General Medical Council's (GMC) document 'Tomorrow's doctors' since the 1970s **[1]**. Most of the students, in the medical sector, cannot able to get the relevance of probability and statistics in their future use as a doctor & as a result find motivation hard. Students might be bound to see the value in the value of such education to their future professions assuming there is proof of how practicing doctors use and value these skills in their work **[2]**.

In 1991 Altman and Bland thought about why doctors had to be aware of statistics and proposed that perusing and deciphering research was the primary reason, alongside understanding pharmaceutical organization literature & diagnostic test and doing one's exploration regardless of whether just for once **[3]**. From the 1990s evidence-based medicine took hold and welcomed an expanded emphasis on basic evaluation abilities, which incorporate some statistical understanding. Morris detailed a shift away from making estimations toward basic evaluation and deciphering insights in the literature, vital ingredients of evidence-based practice **[4]**. Similarly, Palmer proposed that all clinical students should become at least consumers of research and that 21st-century doctors will require ordnance of basic research abilities to survey the Internet-writing of differing characteristics presented by their patients **[5]**.

While writings the perspective of statistics and probability-based research papers, the doctors should understand well about probability and statistics. In 1997, a survey of general practitioners (GPs) U.K. regarding their engagement with evidence-based practice proposed that a variable but sizeable minority did not understand statistical terms used in the evidence sources **[6]**. In different another survey it is also declared that knowledge of & attitudes toward statistics and probability in the framework of EBM practice **[7, 8]**, they have not investigated the features of the doctor's responsibility for which an understanding of probability & statistics would be helpful.

we are all facing certain events in our day-to-day life, and the occurrence of events is part of life. Apart from this, variation is also a part of life, if a researcher wants to find out what is the probability of India may face the fourth wave of the COVID-19 pandemic. How does the researcher deal with this question? Here they follow a probability distribution to deal with such kinds of questions. For a suitable inference and decision making many healthcare problems can be solved using a probability distribution. In developing countries like India, we are facing a lot of events in day to life; here we can apply the concept of probability and random variables to our event. So that it will help us to calculate how likely the event is certain or uncertain. The most popular branch of statistics is probability. It has a wide range of distribution, all most every field the use of probability and statistics is essential. Here we aim to find out the inferences from the data by using probability and statistics in the health care system for its good management.

Random variables

A random variable contains numerical values derived from a random experiment **[15]**. These are the kind of variables that assigns some numeric values to each outcome of an experiment **[16]**. Basically, in probability, a random variable is of two types such as discrete random variable, which can give specific values, and continuous random variables, which can take any values within a certain range. Several diagnoses of a patient is an examples of a discrete random variables. The body mass index of a patient can take any value like whole number as well as fraction. Hence this is the case of a continuous random variable.

- In most of the linear models (Regression analysis), we can use random variables.
- In most of the general linear models (logistic Regression analysis), we can use random variables.
- Random variables are used to estimate the probability of an adverse event occurring by the healthcare analyst and risk analyst.

Definition of Probability

The term probability is difficult to define. In common terms, we can define probability, as the chance of happening or not happening of an event in a random experiment. it is based on the relative frequency of an observed event, observed in past circumstances **[9]**. Whenever we are calculating the probability of any event, the value must lie between 0 to 1. If an event is sure to happen then the probability must be 1, calculating the probability of events is very important in the healthcare system. It enables us to decide conditions of uncertainty with a calculated risk.

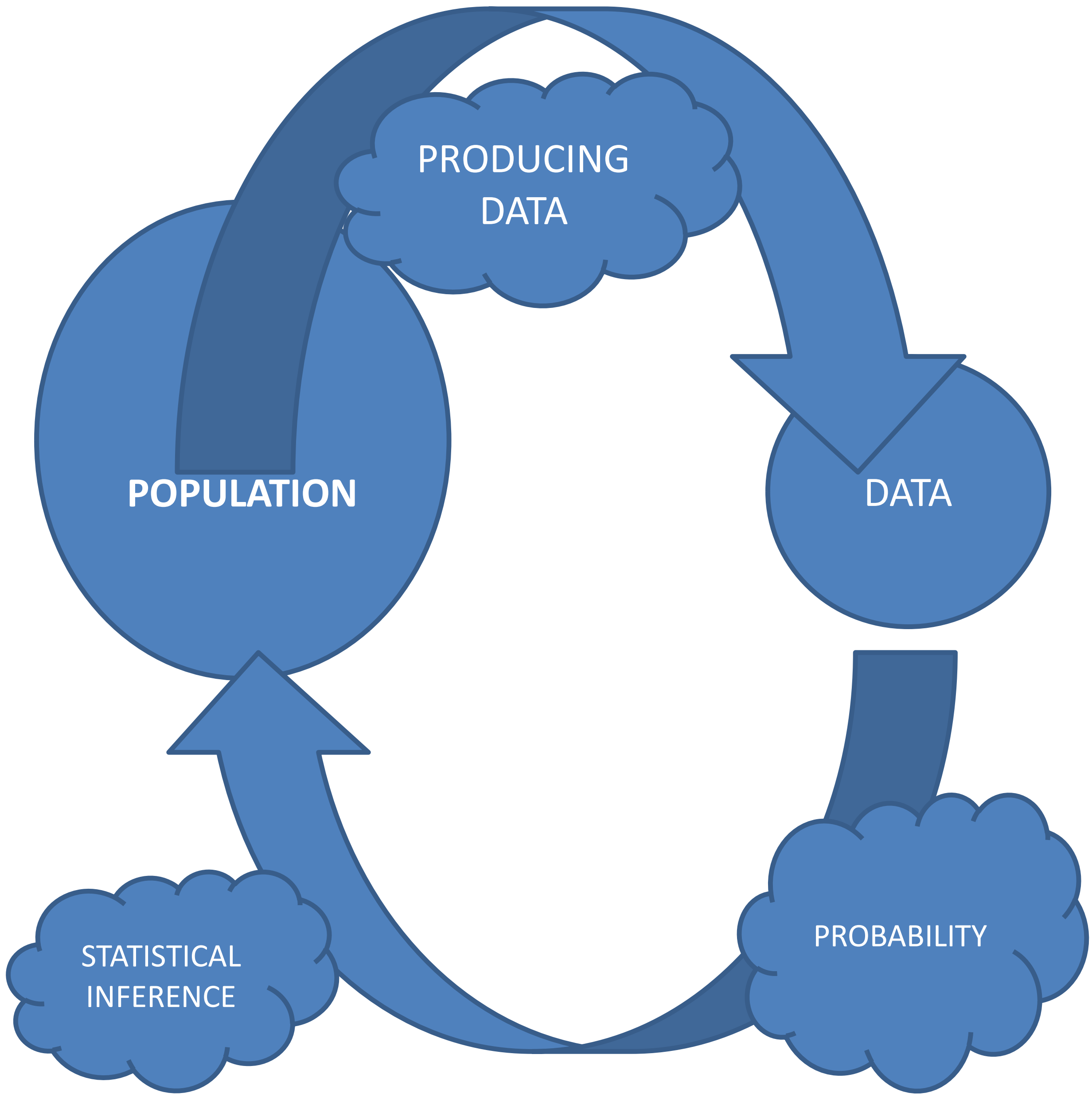

Figure 1: Road map for probability and inference from the data

Binomial distribution

A Very popular example of a probability family is a binomial distribution. When two events are mutually exclusive, then the binomial distribution is useful. This type of distribution has only two possible outcomes. However, binomial distribution may describe also two events that are mutually exclusive but are not equally possible. For example that a newborn baby will contain normal birth weight or low birth weight.

We can define the binomial distribution as below. Let us consider we are conducting "n" number of trials and the probability of success is defined as p and the probability of failure is "q". To find "x" success in "n" trial we can use the binomial probability distribution as below-

P(X=x) = $c_x^n\, p^x q^{n-x}$**……………………..(1)**

Table 1: Sum of all possible binomial probability distribution

Number of successes	Probability p(x)
0	$c_0^n\, p^0 q^{n-0}$
1	$c_1^n\, p^1 q^{n-1}$
2	$c_2^n\, p^2 q^{n-2}$
3	$c_3^n\, p^3 q^{n-3}$
x	$c_x^n\, p^x q^{n-x}$
n	$c_n^n\, p^n q^{n-n}$
Total	1

As it is shown in the **Table 1**, the total sum of all possible probabilities of a distribution is 1.

$$\sum_{x=0}^{n} p(x) = p(0) + p(1) + p(2) + \cdots \ldots\ldots\ldots\ldots\ldots . p(n)$$

$$= c_0^n \; p^0 q^{n-0} + c_1^n \; p^1 q^{n-1} + c_2^n \; p^2 q^{n-2} + \cdots \; c_n^n\, p^n q^{n-n}$$

$$= (p+q)^n = 1$$

……………………………………………(2)

Let us consider a simple example to understand the binomial distribution. The probability of success of a doctor in a certain operation is 80%. if the doctor is attending 25 operations in a certain hospital, what is the probability he will succeed in exactly 20 operations? For this inference, we shall use the concept of a binomial probability distribution.

Here Number of trial is 25.i.e n=25.probability of success is 0.8.probability of failure is 1-p=0.2

We need x=20. Hence using the binomial probability distribution we can calculate as below

P(x=20) = $c_x^n \; p^x q^{n-x} = c_{20}^{25}\; 0.8^{20}\, 0.2^{25-20}$

This can be written as P(x=20) = $c_{20}^{25}\; 0.8^{20}\, 0.2^{5}$ = 0.1960

To find out such calculations we can use the r programming language or any online software or we may use excel to find out the binomial probability distribution. In excel we can use the function BINOMDIST (Numbers, trial, Probability, cumulative).

In another case, if in a hospital the data shows that the patients suffering from COVID-19, 75% recover from it. What will be the probability if we select 16 random patients out of which 13 will recover from covid 19?

Here we shall apply the BINOMDIST function. The probability that 13 will recover using density distribution at all points. Now let us execute the above example in excel.

Here x= 13 n=16

P(x=13) = $c_0^n\, p^0 q^{n-0}$ = $c_{13}^{16}\, 0.75^{13} 0.25^{16-13}$ **=0.2078**

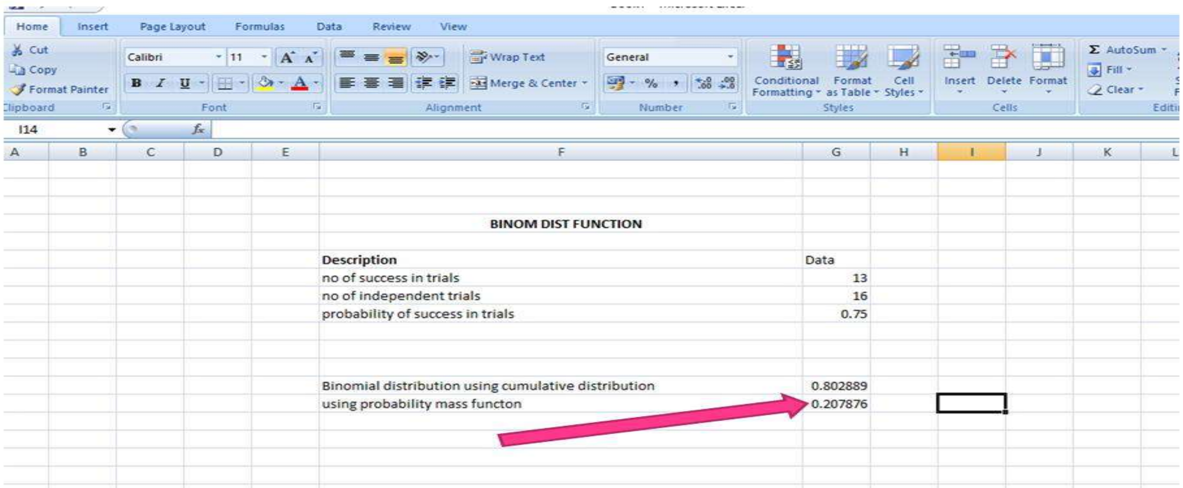

Figure 2: Binomial probability distribution using excel function

5. Poisson distribution

Poisson distribution is another important distribution in probability. It is a discrete distribution. In many situations the data, we are using in the healthcare system follows the Poisson distribution. It is useful to describe the probability that a given event can happen within a given period **[10, 11]**. The characteristics of the Poisson are given below.

I) all the events are independent from each other in the random experiment.

II) The event may present from 0 to ∞ within a given interval.

(III) When the period of observation is longer, the probability of an event to happen increases.

The Poisson distribution follows the following formula

F(x) = $\frac{e^{-\lambda}\lambda^{x}}{x!}$ **.. (3)**

Where λ is the mean of the Poisson distribution. x is no of success and e is the mathematical constant e = 2.718

Now Let us discuss a case to understand the concept of Poisson distribution in healthcare. In a particular hospital, an average of 13 new cases of oral cancer are diagnosed each year. Suppose the annual occurrence of oral cancer is distributed by Poisson distribution .suppose we need to find out the probability that in a given year the number of newly diagnosed cases of oral cancer will be

1) exactly 10 II) at least 8 iii) between 9 and 11, inclusive.

Here comes the application of Poisson distribution in real life. Now to find out the probability of the above question we need to follow the formula of Poisson distribution.

We have e = 2.718 λ = 13

I) P(x=10) = $\frac{e^{-\lambda}\lambda^{x}}{x!}$ = $\frac{2.718^{-13}13^{10}}{10!}$ = 0.086

It means there is a chance that in a given year, exactly 10 newly diagnosed cases of oral cancer will be 8% in that hospital.

II) P(x ≥8) = 1-[p(x=0)+ p(x=1)+ p(x=3)+ p(x=4)+ p(x=5)+ p(x=6)+ p(x=7)]

P(x=0) = $\frac{2.718^{-13}13^{0}}{0!}$ =0.0000022

P(x=1) = $\frac{2.718^{-13}13^{1}}{1!}$ =0.000029

P(x=2) = $\frac{2.718^{-13}13^{2}}{2!}$ =0.00019

P(x=3) = $\frac{2.718^{-13}13^{3}}{3!}$ =0.00083

P(x=4) = $\frac{2.718^{-13}13^{4}}{4!}$ =0.0027

P(x=5) = $\frac{2.718^{-13}13^{5}}{5!}$ =0.007

P(x=6) = $\frac{2.718^{-13}13^{6}}{6!}$ =0.015

P(x=7) = $\frac{2.718^{-13}13^{7}}{7!}$ =0.028

Total = 0.054

Hence the required probability = 1-0.054=0.946 (94%)

It means the probability that in a given year at least 8 numbers of newly diagnosed cases of oral cancer will be 94% in that hospital.

III) To find the probability between 9 and 11, inclusive we need to write

P(X=9)+P(X=10)+P(X=11)

$\frac{2.718^{-13}13^{9}}{9!} + \frac{2.718^{-13}13^{10}}{10!} + \frac{2.718^{-13}13^{11}}{11!} =$ 0.066+0.086+0.102=0.254(25%)

It means the probability that in a given year newly diagnosed cases of oral cancer between 9 to 11 will be 25% in that hospital.

6. NORMAL DISTRIBUTION

The normal distribution is the most frequently used distribution in probability **[13, 14]**. Most of the real-life applications correlate with the normal distribution. it is one of the theoretical distributions. Many natural physical and biological phenomena which exhibit frequency distribution, also closely resemble the normal distribution **[12]**. In most, biological analyses, values are often distributed following the normal distribution. it is a continuous probability distribution, and the density function is defined as

$$\mathbf{F(x)} = \frac{1}{6\sqrt{2\Pi}} \cdot e^{\frac{-(x-\mu)^2}{26^2}} \quad -\infty<\mathbf{x}<\infty \text{ ....(4)}$$

X = Numeric values of random variables

μ = Mean

Ϭ = standard deviation of a normal random variable

e = 2.718 Π = 3.1416 $\sqrt{2\Pi}$ =2.5066

The normal distribution is represented by a smooth curve .and the curve is perfectly symmetrical and continuous. it is a bell-shaped curve and extends on both sides indefinitely. The normal distribution is a continuous probability distribution hence the area under the curve is 1. as we know the total probability is always 1.

For μ=0 and Ϭ=1 we can define the curve standardized normal curve. It can be defined as

$$\mathbf{Z} = \frac{(x-\mu)}{6} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\mathbf{(5)}$$

$$\mathbf{f(z)} = \frac{1}{\sqrt{2\Pi}}.e^{\frac{-z^2}{2}} \quad -\infty<x<\infty \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\mathbf{(6)}$$

Let's consider one example to illustrate normal distribution. Suppose the age at which the children suffer from a particular disease is normally distributed with a mean of 11.5 years and a standard deviation of 3 years. if a child has come down with the disease then what is the probability that the child's age will be under 12?

Here given x=12 mean(μ) = 11.5, sd(Ϭ) = 3

Hence converting it into standard normal distribution $Z=\frac{(x-\mu)}{6} = \frac{(12-11.5)}{3} = 0.167$

The area between z = 0 and z= 0.167 is 0.0675(Using table).Now the total area will be 0.5+0.0675=0.5675. So we can say there is a chance of 56% ,if a child come down with the disease age will be under 12.

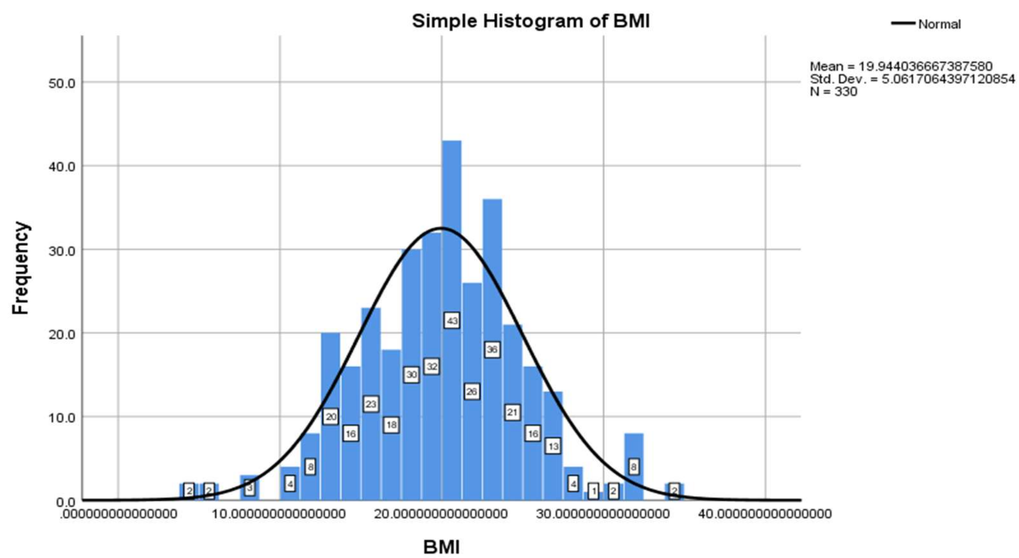

Figure 3: Distribution of body mass index

7. CONCLUSION

This paper leads us to understand that most of the data in healthcare follow a probability distribution, either discrete distribution or continuous distribution. The concept of p-value is based on probability, which we are using in most healthcare analyses. P value is the probability of being wrong when considering that a difference exists. Many health care models are related to a probability distribution. The cream of the study will be a great contribution of probability to healthcare analysis and healthcare management.

REFERENCE

[1] General Medical Council. Tomorrow's Doctors: Recommendations on Undergraduate Medical Education. General Medical Council: London, 2003.

[2] Astin J, Jenkins T, Moore L. Medical students' perspective on the teaching of medical statistics in the undergraduate medical curriculum. Statistics in Medicine 2002; 21(7):1003–1006.

[3] Altman DG, Bland JM. Improving doctors' understanding of statistics. Journal of the Royal Statistical Society, Series A 1991; 154(2):223–267.

[4] Morris RW. Does EBM offer the best opportunity yet for teaching medical statistics? Statistics in Medicine 2002; 21(7):969–977.

[5] Palmer CR. Discussion: teaching hypothesis tests: time for significant change? Statistics in Medicine 2002; 21(7):995–999.

[6] McColl A, Smith H, White P, Field J. General practitioners' perceptions of the route to evidence based medicine: a questionnaire survey. British Medical Journal 1998; 316:361–365.

[7] Oliveri RS, Gluud C, Wille-Jorgensen PA. Hospital doctors' self-rated skills in and use of evidence-based medicine—a questionnaire survey. Journal of Evaluation in Clinical Practice 2004; 10(2):219–226.

[8] Windish DM, Huot SJ, Green ML. Medicine residents' understanding of the biostatistics and results in the medical literature. Journal of the American Medical Association 2007; 298(9):1010–1022.

[9] Daniel WW. eds. Biostatistics: a foundation for analysis in the health sciences. New York: John Wiley & Sons, 1995.

[10] Kolmogorov AN. eds. Foundations of Theory of Probability. Oxford: Chelsea Publishing, 1950.

[11] Lim E. Basic statistics (the fundamental concepts). J Thorac Dis 2014;6:1875-8

[12] Standard Normal Distribution Table. Available online: http://www.mathsisfun.com/data/standard-normaldistribution-table.html

[13] Pauker SG,Kassirer JP. The threshold approach to clinical decision making. N Engl J Med. 1980; 302 (20):1109–

17. https://doi.org/10.1056/NEJM198005153022003 PMID: 7366635.

[14] Evidence-based health care: a new approach to teachingthe practice of health care. Evidence-Based Medicine Working Group. J Dent Educ.1994;58(8):648–53. PMID:7930000

[15] Eisenberg JM, HersheyJC.Derived thresholds. Determining the diagnostic probabilities at which clinicians initiate testing and treatment. Med Decis Making. 1983; 3(2):155–68. https://doi.org/10.1177/0272989X8300300203PMID:6415358

[16] Dempster, A. P. (1968). A Generalization of Bayesian Inference. Journal of the Royal Statistical Society: Series B (Methodological), 30, 205–232. https://doi.org/10.1111/j.2517-6161.1968.tb00722.x