Online ISSN : 2249-4626
Print ISSN : 0975-5896
DOI : 10.17406/GJSFR

Global Journal

of Science Frontier Research: F

Mathematics and Decision Science

Highlights

Forecasting of Fog

Point Theorem in Economics

Fuzzy Interference Systems

Solutions of Fermat's Equation

Discovering Thoughts, Inventing Future

VOLUME 21 ISSUE 5 VERSION 1.0

Global Journal of Science Frontier Research: F
Mathematics & Decision Sciences

Global Journal of Science Frontier Research: F
Mathematics & Decision Sciences

Volume 21 Issue 5 (Ver. 1.0)

Open Association of Research Society

Incorporation No.: 0423089
License No.: 42125/022010/1186
Registration No.: 430374
Import-Export Code: 1109007027
Employer Identification Number (EIN):
USA Tax ID: 98-0673427

Global Journals Inc.

*(A Delaware USA Incorporation with "Good Standing"; **Reg. Number: 0423089**)*
Sponsors: *Open Association of Research Society*
Open Scientific Standards

Publisher's Headquarters office

Global Journals® Headquarters
945th Concord Streets,
Framingham Massachusetts Pin: 01701,
United States of America
USA Toll Free: +001-888-839-7392
USA Toll Free Fax: +001-888-839-7392

Offset Typesetting

Global Journals Incorporated
2nd, Lansdowne, Lansdowne Rd., Croydon-Surrey,
Pin: CR9 2ER, United Kingdom

Packaging & Continental Dispatching

Global Journals Pvt Ltd
E-3130 Sudama Nagar, Near Gopur Square,
Indore, M.P., Pin:452009, India

Find a correspondence nodal officer near you

To find nodal officer of your country, please email us at *local@globaljournals.org*

eContacts

Press Inquiries: *press@globaljournals.org*
Investor Inquiries: *investors@globaljournals.org*
Technical Support: *technology@globaljournals.org*
Media & Releases: *media@globaljournals.org*

Pricing (Excluding Air Parcel Charges):

Yearly Subscription (Personal & Institutional)
250 USD (B/W) & 350 USD (Color)

Editorial Board

Global Journal of Science Frontier Research

Contents of the Issue

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 21 Issue 5 Version 1.0 Year 2021
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Parametric Solutions of Fermat's Equation

By Sudhangshu B. Karmakar

Abstract- The values of the variables of the equation, $a^n + b^n = c^n$, are first obtained in terms of two parameters g and h. By substituting these values in the equation it is verified that these parametric solutions as obtained indeed satisfy the Fermat's Equation.

GJSFR-F Classification: MSC 2010: 11D41

PARAMETRICSOLUTIONSOFFERMATSEQUATION

Strictly as per the compliance and regulations of:

Parametric Solutions of Fermat's Equation

Sudhangshu B. Karmakar

Abstract- The values of the variables of the equation, $a^n + b^n = c^n$, are first obtained in terms of two parameters g and h. By substituting these values in the equation it is verified that these parametric solutions as obtained indeed satisfy the Fermat's Equation.

I. Introduction

Ever since the French mathematician Pierre de Fermat[1] stated the conjecture in 1630 that the equation, $a^n + b^n = c^n$, cannot have any solutions if a, b, c are non-zero integers each > 0 and n is an integer > 2, the equation has been a subject of intense and often heated discussions amongst mathematicians and aspirants alike. This conjecture is known as Fermat's Last Theorem. The fact that Fermat claimed to have a proof but never wrote it down has put the researchers in a quandary since nobody has yet been able to duplicate the proof the way Fermat originally claimed. Perhaps, the strong appeal of the problem is the simplicity and elegance of its statement contrasted with apparent hopelessness [2] of an elementary way to establish it. Finally, in around 1995 Wiles [3, 4] offered a proof of the Theorem. His paper incorporates by reference [5 ,6] a vastly larger body of mathematical work developed over the last several decades. And it requires an extraordinary arsenal [7] of mathematical tools to understand Wales's complex and very lengthy proof. Consequently, the quest for a simple and short proof continues. In this paper, based on elementary principles, simple parametric solutions of the equation are presented.

II. The Theorem

Fermat asserted [6,7,8] that if a, b, c are nonzero positive integers and n is an integer > 2 then there is no solution of the equation (1).

$$a^n + b^n = c^n \tag{1}$$

This paper offers a simple proof of the theorem based on elementary principles.

III. Simplification of the Theorem

It is enough to prove [1,7] the theorem when the variables are relatively prime integers designated as (a, b, c) = 1, 2|b and the exponent is a prime k > 3.

Ref

1. Ribet, K. A., Hayes, B., Fermat's Last Theorem and Modern Arithmetic, Amer. Sci. 82,1994,144-156.

Author: *Shirley Parkway, Piscataway, USA. e-mail: math75321@gmail.com*

IV. Strategy of the Proof

Lemma-1 2

To be candidates for integer solutions of (1) the variables a, b, c must form the three sides of an acute angled triangle ABC under the given conditions.

$$a^k + b^k = c^k \tag{1}$$

Conditions: (a, b, c) = 1, c > a > b, 2|b, prime k > 3.

Proof:

Case-1: $a + b <= c$ then $a^k + b^k + M < c^k$, $M > 0$. So $a^k + b^k = c^k$ is impossible.
Case-2: $a + b > c$. Therefore, a, b, c must form the three sides of the triangle ABC.

If ABC is a nonacute triangle then $a^2 + b^2 <= c^2$ then $c^{k-2}\, a^2 + c^{k-2}\, b^2 <= c^k$
Since $c > a > b$. Obviously, $a^k + b^k < c^k$. Hence $a^k + b^k\, c^k$ is impossible.

Therefore, ABC must be an acute angled triangle.

This proves Lemma-1

Proof of the Theorem.

To prove the theorem an auxiliary equation (1.1) is introduced:

$$x^{2k} + y^{2k} = z^{2k} \tag{1.1}$$

By comparing (1) and (1.1) it is seen that $a = x^2$, $b = y^2$ and $c = z^2$
Therefore, by obtaining solutions of (1.1) solutions of (1) can be obtained.

The variables are first assumed to be integers each > 0.
Under these assumptions parametric solutions of (1.1) are obtained.
To prove the theorem the variables are first to be nonzero integers and a parametric solution of the equation is obtained. And then it is argued that these parametric solutions cannot lead to integer solutions of the equation.

Parametric Solutions of the Equation:

V. Hypothesis-1

There exist integers g and h such that (2) and (3) are satisfied [10,p-536] for integers such that $X = x^2, Y = y^2$, (x, y) = 1 (g, h) = 1, 2|h.

$$X + iY = (g + ih)^k \quad 2) \qquad X - iY = (g - ih)^k \tag{3}$$

By multiplying the corresponding sides of (2) and (3) one gets (4) where tan H = h/g, $0 < H < \pi/2$.

$$X^2 + Y^2 = (g^2 + h^2)^k\, [(\cos kH)^2 + (\sin kH)^2] \tag{4}$$

Notes

Since $(\cos kH)^2 + (\sin kH)^2 = 1$ one gets (5).

$$X^2 + Y^2 = (g^2 + h^2)^k \tag{5}$$

By comparing (4) and (5) (6) is obtained.

$$g^2 + h^2 = z^2 \tag{6}$$

Therefore, the variables g, h, z are the three sides of a right triangle ZGH where z is the hypotenuse. Since all the sides and the area of the triangle are rational it is a rational[11] right triangle. Therefore, one gets (7) and (8).

$$x = z(\cos kH)^{1/k} \tag{7}$$

$$y = z(\sin kH)^{1/k} \tag{8}$$

It is verified that (7) and (8) are indeed the parametric solutions of (1.1)

Now by noting that $a = x^2$, $b = y^2$, $c = z^2$ (7.1) and (8.1) are obtained.

$$a = c(\cos kH)^{2/k} \quad (7.1) \qquad b = c(\sin kH)^{2/k} \quad (8.1)$$

By substituting these values of a, b in (1) it is verified that a, b, c as given by (7.1) and (8.1) are indeed the parametric solutions of (1).

Next challenge is to prove that given c is an integer > 0, (a, b) = 1 is impossible in (7.1) and (8.1)

References Références Referencias

1. Ribet, K. A., Hayes, B., Fermat's Last Theorem and Modern Arithmetic, Amer. Sci. 82,1994,144-156.
2. Mozzochi, C.J., The Fermat Proof, Trafford, 2004.
3. Wiles, A., Modular elliptic curves and Fermat's Last Theorem, Ann. of Math.141,1995,443-551.
4. Faltings G., The Proof of Fermat's Last Theorem, Notices of the AMS, July 1995.
5. Hellegouarch, Y., Invitation to the Mathematics of Fermat – Wiles, Academic Press, 2004.
6. Van der Poorten, A., Notes on Fermat's Last Theorem, John Wiley, New York, 1996.
7. Ribenboim, P., 13 Lectures on Fermat's Last Theorem, Springer-Verlag, New York,1979.
8. Dickson, L.E., History of the theory of numbers, Chelsea, New York, 1971.
9. Granville, A., Monagan, M.B., The first case of Fermat's Last Theorem is true for all prime exponents up to 714, 591, 416, 091, 389, Trans. Amer. Math Soc. 306(1988)No 1, 329-359.

Ref

11. Lehmer, D. N., Rational Triangles, The Annals of Math.Vol.1, No.1/4(1899-1900), 97-10.

10. Darmon, H., Granville, A., On the Equation $z^m = F(x,y)$ and $Ax^p + By^p = Cz^r$, Bull. London Math Soc. 27(1995), 513-543.
11. Lehmer, D. N., Rational Triangles, The Annals of Math.Vol.1, No.1/4(1899-1900), 97-10.

Notes

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 21 Issue 5 Version 1.0 Year 2021
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Forecasting of Fog by using Fuzzy Interference Systems

By Md. Anisur Rahman
Islamic University

Abstract- In this study the method of fuzzy interference systems are used to evaluate fog forecasting. Input variables used in this study included the situation of fog parameters dew point, dew point spread, rate of change of dew point spread wind speed and sky coverage. The membership functions are generally trapezoids, although simpler functions such as triangles and rectangles and even delta functions are often used. The degrees of membership of the system inputs are also examined. The strength of a rule has been derived from the corresponding degrees of membership of the system inputs. Only 16 rules of the entire set of 144 would have non-zero values, or strengths. The techniques indicate some preliminary results using real data. The final system output is then calculated as the weighted average of the centroid of each membership function with the area of the enclosed set used as the weighting factor. These centroids and weighting areas of the current values were calculated to find the forecasting of fog. The result of this study would hopefully help the planners and program managers to take necessary actions and to development air, marine, and road traffic etc.

Keywords: *fuzzy logic, membership function, parameter, forecast, fog.*

GJSFR-F Classification: *MSC 2010: 03B52*

FORECASTING OF FOG BY USING FUZZY INTERFERENCE SYSTEMS

Forecasting of Fog by using Fuzzy Interference Systems

Md. Anisur Rahman

Abstract- In this study the method of fuzzy interference systems are used to evaluate fog forecasting. Input variables used in this study included the situation of fog parameters dew point, dew point spread, rate of change of dew point spread wind speed and sky coverage. The membership functions are generally trapezoids, although simpler functions such as triangles and rectangles and even delta functions are often used. The degrees of membership of the system inputs are also examined. The strength of a rule has been derived from the corresponding degrees of membership of the system inputs. Only 16 rules of the entire set of 144 would have non-zero values, or strengths. The techniques indicate some preliminary results using real data. The final system output is then calculated as the weighted average of the centroid of each membership function with the area of the enclosed set used as the weighting factor. These centroids and weighting areas of the current values were calculated to find the forecasting of fog. The result of this study would hopefully help the planners and program managers to take necessary actions and to development air, marine, and road traffic etc.

Keywords: *fuzzy logic, membership function, parameter, forecast, fog.*

I. Introduction

Fog forecasting is one of the most important and demanding operational responsibilities carried out by meteorological services worldwide. It is a difficult process that includes numerous specialized technological fields. Fog is a type of weather phenomenon that reduces the atmospheric horizontal visibility to below 1 km due to the suspending of ice or water droplets in the atmosphere near the surface [1]. Fog is an observable aerosol consisting of little water droplets or ice crystals floating in the air at or near the Earth's surface. It can be considered a type of low-lying cloud generally similar to stratus, and is seriously influenced by nearby bodies of water, topography and wind conditions. In turn, fog has affected many human activities, such as shipping, travel, aircraft communication and warfare [2]. The effect of fog on human life was recognized in the early ages of mankind but its impact has significantly increased during recent decades due to increasing air, marine, and road traffic. In fact, the financial and human losses connected to fog and low visibility became similar to the losses from other weather events, e.g., tornadoes or, in some situations, even hurricanes. Moreover, when the fog contains heavy pollutants, it is also harmful to human health [3].

Recently, many studies revealed that when stern fog formed, it generally showed obvious volatile features, including sudden surge of fog drop number density, obvious increase in the fog drop scale and water content, sharp plunge of visibility from several hundred meters to below 50 meters, and fast changing from heavy fog to severe fog

Ref

1. T. S. Glickman, Glossary of Meteorology, 2nd edition, American Meteorological society, 2000: 855.

Author: *Associate Professor, Department of Mathematics, Islamic University, Kushtia-7003, Bangladesh.*
e-mail: anismathiu@yahoo.com

within a very short time period (about 30 min) [4]. With the improvement of temporal resolution of droplet spectrometer, [5] analyzed microphysics processes and macroscopic situation of fog droplet spectrum widening. He pointed out that the former stage of fog droplet spectrum widening mainly occupied nucleation and condensation processes, while the latter stage mainly involved coagulation and concentration processes. Turbulence not only plays a significant role in heat, momentum, and moisture vertical transfers but also is the necessary condition for coagulation and increase of fog droplet. Radioactive cooling alone cannot produce large fog droplets, while greater super saturation and turbulence can promote the formation of large droplets [6]. Although fog and haze are different weather phenomena, their formations have similarity. They all form under breeze, high humidity, and static stability condition [7]. The haze aerosols can be transformed to fog droplets under certain conditions [8]. With the increasingly heavy traffic, there has been a recent increase in demand for accurate fog forecasting. However, fog forecasting is currently still challenging. Current performance of fog forecasting is much lower than that of precipitation forecasting from the same operational prediction systems at the national centers for environmental prediction [9]. Beside model resolutions, fog forecasting has also shown to be highly sensitive to initial conditions [10]. With the rapid evolving technologies in the field of meteorology, it is desirable to merge the experience of many forecasters with algorithms that may aid in difficult forecasting situations [11]. One such method is called fuzzy logic. Fuzzy logic is a simple yet powerful problem solving technique with widespread capabilities. It is currently used in the fields of business, systems control, electronics and traffic engineering. The technique can be used to generate solutions to problems based on vague, ambiguous, qualitative, incomplete or imprecise information [12].

There are many weather forecasting difficulties that fit the previous description where precise, analytical solutions do not exist. How precise the culturally built original climate knowledge of extreme climatic events is, and how acquiescent it is to fuzzy logic [13]. Fuzzy set theory is a tool for modeling the kind of uncertainty associated with ambiguity, feeling and/or with a lack of information regarding particular element of the problem in hand [14]. Developed fuzzy logic model is made up of two functional components; the knowledge base and the fuzzy reasoning or decision making unit [15]. Fuzzy inference is the actual procedure of mapping with a given set of input and output through a set of fuzzy systems [16]. Two operations were performed on the fuzzy logic model; the fuzzification operation and defuzzification operation. Fuzzy logic based forecasting method by using the Mamdani fuzzy inference system may be successively used for different environmental problem assessment to moderate unexpected meteorological problems [17]. Hence the study is based on the application of Mamdani fuzzy inference system in one of the meteorological procedure, which is fog in order to get forecasting of the region. The aim of this study is to develop a fuzzified model for fog forecast and explores the use of fuzzy logic in fog forecasting.

II. Methodology

Fuzzy logic is a methodology to delicacy uncertainty in systems whose variables are rather qualitative, imprecise or ambiguous. Over the past few years, this methodology has been increasingly adopted as a constructive technique in a variety of environmental issues but has restricted applications in weather prediction. The appropriateness of fog occurrence for conduct with fuzzy logic has been tested with hopeful results [18, 19]. More recently, a fuzzy logic is system for the analysis and prediction of cloud ceiling and visibility [20]. Also reports on the use of such a

Ref

5. Z. Li, J. Yang, C. Shi and M. Pu, Urbanization effects on fog in china: field research and modeling, *Pure and Applied Geophysics*, 2012, 169(5):927–939.

forecasting system to forecast ceiling and visibility and concludes that the fuzzy system outperforms persistence-based forecasts [21]. Fuzzy sets are the collection of objects with the same properties, and in crisp sets the objects either belong to the set or do not. In practice, the characteristic value for an object belonging to the considered set is coded as 1 and if it is outside the set then the coding is 0. In crisp sets, there is no ambiguity or vagueness about each object belongs to the considered set. So uncertainty always arises concerning the assessment of membership values 0 or 1. In order to assuage such situations generalized the crisp set membership degree as having any value continuously between 0 and 1 [22]. An object with membership function 1 belongs to the set with no doubt and those with 0 membership functions again absolutely do not belong to the set, but objects with transitional membership functions partially belong to the same set. Fuzzy set theory has been examined in hydrological modeling such as fog forecasting, inflow prediction and reservoir operation. Several reservoir operation techniques based on fuzzy systems and fuzzy control systems for reservoir operations can be found in the literature along with the use of genetic algorithms. These models include fuzzy optimization techniques, fuzzy rule based systems and hybrid fuzzy systems. In these models, fuzzy rules were created using expert knowledge or observed data. In fuzzy logic, values of variables are expressed by linguistic terms, the relationship is defined in terms of IF-THEN rules and the outputs are also fuzzy subsets which can be made crisp using defuzzification techniques [23]. Finally, it is expected that the findings of the study will support fog observing and the outcomes will demonstrate weather forecasting in the future.

Ref

21. B. K. Hansen, Analog forecasting of ceiling and visibility using fuzzy sets, 2nd Conference on Artificial Intelligence, *American Meteorological Society*, 2000, 1–7.

III. Fuzzy Set

Let X be a universal set. Then A is called a (fuzzy) subset of X if A is a set of ordered pairs.

$$A = \{(x, \mu_A(x)) : x \in X, \mu_A(x) \in [0,1]\}$$

where μ_A is the membership function of A and $\mu_A(x)$ is the grade of the membership of x in A.

The linguistic expression for the variables and their membership functions are evaluated from the following triangular membership functions and it is defined by a lower limit a, an upper limit b and a value m, where $a \le m \le b$. The value of the membership function $\mu(x)$, ranges from 0 to 1, with 0 denoting no membership, 1 for full membership and values in between has partial membership as shown in figure 1.

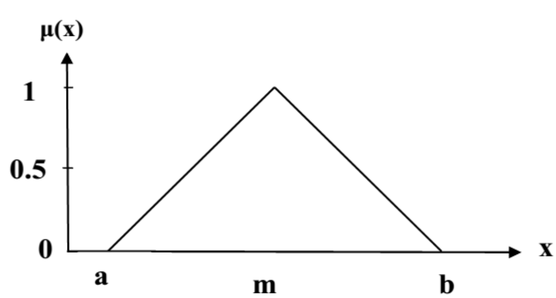

Figure 1: Basic graph of Fuzzy membership function

The methodology employed is referred to as the centre-of-gravity method of trapezoidal diagram in which the area of all applicable output sets, limited in height by the applied rule strength, are evaluated and a weighted mean value of the output is calculated.

Therefore the fog forecasting $= \frac{\sum C_i A_i}{\sum A_i}$, where $C_i =$ weighted average of the centroid of each membership function and $A_i =$ area of the trapezium enclosed set used as the weighting factor.

Notes

IV. Fuzzy Interference Systems

The fuzzy inference system is an important part of the fuzzy logic system. Decision-making is the main work of it performs. It makes use of the rules of IF and THEN along with the AND and OR connectors which represent some essential rules for decision making. It is a framework which depicts the actual procedure of converting an input into an output using fuzzy logic. Fuzzification, defuzzification, membership function are the building blocks of fuzzy inference system. Employing fuzzy IF-THEN rules to express input–output relationships and model the qualitative inputs and logic procedure for creating the output. The fuzzy inference systems include a set of precursor and consequential fuzzy membership functions as well as a set of Fuzzy IF–THEN rules which considered a firm basis for developing the core of any system which strength be used for making decisions in vague and inaccurate situations. A fuzzy inference system is impersonating human intelligence for quantify the ambiguity in real world situation through proper modeling of the fuzziness of real world data by suitable rule bases. The IF portion of a rule refers to the degree of membership in one of the fuzzy sets. The THEN portion refers to the consequence, or the associated system output fuzzy set which is discussed below. One rule could be stated:

IF (dry AND unsaturated AND drying AND light AND cloudy) THEN (low probability of fog).

The rules must be constructed for each system and must rely on the experience of the forecaster. The total number of rules is the product of the number of fuzzy sets characterizes the system.

V. Fuzzy Systems Input

The meteorological observable fact select here to reveal this method is that of the occurrence of fog. Several techniques have been developed and used over the years to assist in the forecasting of the beginning of the fog. Although no analytic method has yet been discovered, the basic physical principles involved in fog forecasting are thought to be well understood. Perspective of the current model of fog forecasting, general terms associated with fuzzy logic will be introduced. In general, a problem to be solved is referred to as a system. System inputs are those substantial variables that are thought to completely determine the solution to the problem, or system outputs. In the current model the system output is the probability of the occurrence of fog within the next (approximately) six hours. System inputs to be used in this study are the situation of fog parameters dew point, dew point spread, rate of change of the dew point spread, wind speed and sky coverage. These system inputs are readily available, and their values are naturally measured when a forecaster decides if or when fog will develop. It is understood that there are other substantial variables that forecasters may judge important in the formation of fog. Fuzzy inference is the actual procedure of mapping

Notes

with a given set of input and output through a set of fuzzy systems. The system inputs are categorized into actually significant domains called fuzzy sets. Fuzzy sets are simply qualitative descriptions of the selected domains of the inputs, each of which is thought to have a specific effect on the output. Table 1 show the fog parameters which used in this study.

Table 1: Situation of fog parameters

Dew Point	Dew Point Spread (DPS)	Rate of change of the DPS	Wind Speed	Sky
Dry	Very Saturated	Drying	Light	Cloudy
Moderate	Saturated	Saturating	Excellent	Clear
Moist	Unsaturated		Strong	
Very Moist				

The interpretation of the qualitative description of these fuzzy sets, and the substantial effect on the probability of occurrence of fog should be apparent. Dew point is closely associated to relative humidity, which is the ratio of the pressure of water vapor in a parcel of air relative to the saturation pressure of water vapor in that same parcel of air at a specific temperature. Relative humidity is expressed as a percentage. The value of the current dew point is categorized into one of four fuzzy sets: dry, moderate, moist and very moist. Similarly, the value of the dew point spread is described as unsaturated, saturated or very saturated. The rate of change of the dew point spread (ﬂC/hour) is only characterized by one of two categories, whether the atmosphere is drying (a positive rate) or shows a saturating trend (a negative rate). The wind speed effect upon the probability of fog formation is described as either strong or excellent or light, whose are unfavorable conditions, or being in the range that is favorable for fog formation. Similarly the sky conditions are classified into either cloudy or clear. The determination of the fuzzy sets is derived solely from experience. A developer of a fuzzy system must sensibly decide these categories such that logical systems outputs will be obtained. The fuzzy sets are quantitatively defined by membership functions. These functions are typically very simple functions that cover a specified domain of the value of the system input. The functions are generally trapezoids, although simpler functions such as triangles and rectangles and even delta functions are often used. The fuzzy sets of the fog parameters in this study are shown in Figures 2 through 6.

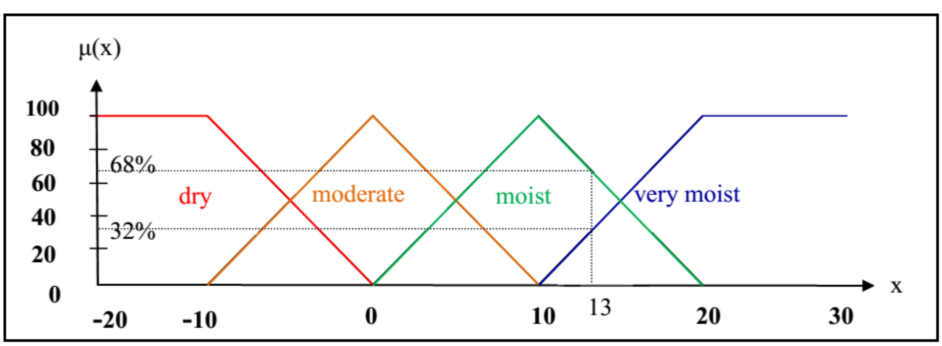

Figure 2: Graph of the membership function of dew point and corresponding fuzzy levels (dry, moderate, moist, very moist)

Global Journal of Science Frontier Research (F) Volume XXI Issue V Version I Year 2021

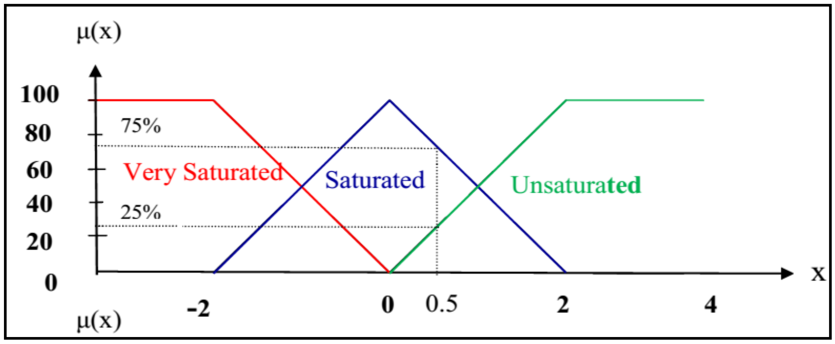

Figure 3: Graph of the membership function of dew point spread and corresponding fuzzy levels (very saturated, saturated, unsaturated)

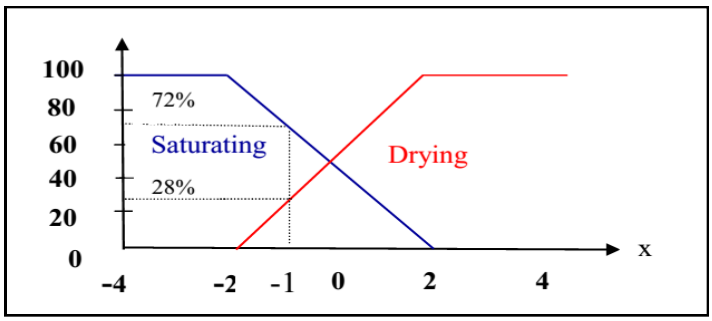

Figure 4: Graph of the membership function of the rate of change dew point spread and corresponding fuzzy levels (saturating and drying)

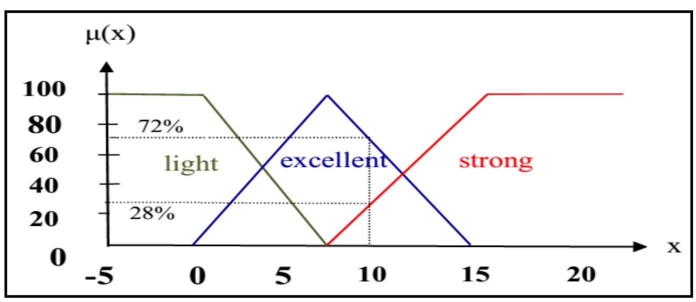

Figure 5: Graph of the membership function of wind speed and corresponding fuzzy levels (light, excellent and strong)

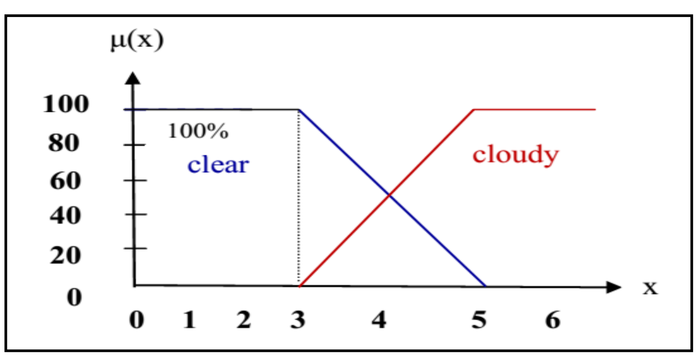

Figure 6: Graph of the membership function of sky condition and corresponding fuzzy levels (clear, cloudy)

Notes

Notes

Each value of system input will belong to at least one fuzzy set and more than one fuzzy set. This is possible because during construction the adjacent fuzzy sets are made to overlap. A rule of thumb is to ensure that the sets overlap by approximately 25 percent. This is the foundation of the technique of fuzzy logic. Dew point value as indicated of 13°C is seen to be both a member of the moist and the very moist sets in Figure 2. The membership function associated with each fuzzy set provides a quantitative value of the degree of membership of the input into each fuzzy set. In Figure 2 a value of 13ﬂC is a 32% member of the moist set, a 68% member of the very moist set. Otherwise dew point spread value surround 0.5°C is seen to be both a member of the saturated and unsaturated sets. In Figure 3, the value of 0.5°C is 75% member of the saturated set and 25% member of the unsaturated set. Whereas the rate of change of dew point spread value contains -1°C/hour is a show to be both a member of the saturating and drying sets. In Figure 4, the value of -1°C/hour is 72% member of the saturating set and 28% member of the drying set. Similarly, the wind speed value indicated 10 kts is a demonstrated to be a member of the light, excellent and strong sets. In Figure 5, the value of 10 kts is 72% member of the excellent set and 28% member of the strong set. The sky condition value indicated 3 oktas is to be a member of the clear and cloudy sets. In Figure 6, the value of 3 oktas is 100% member of the clear set and 0% member of the cloudy set.

The degrees of membership of the system inputs are evaluated as discussed above. The determination of the system outputs (probability of fog formation) follows from the evaluation of a set of predefined rules. The strength of a rule is derived from the corresponding degrees of membership of the system inputs. Since an input can be member of multiple fuzzy sets, then different rules involving these sets can be applied. The higher degrees of membership result in corresponding rules which have more strength in the final estimated procedure. In other words, the number of rules equals all possible permutations of categorized system inputs. Here there are four sets associated with the dew point, three with the dew point spread, two with the rate of change of the spread, three with the wind speed and two with the sky condition. The total number of rules that completely define the set then is 4x3x2x3x2=144. It is easy to see how the number of rules rapidly expands with each system input and related fuzzy sets. The strength of a rule is the value of its least true antecedent, or IF portion, which is simply the degree of membership of each system input in the corresponding fuzzy sets. More than one rule can lead to the same consequence. In this case, the rule with the highest strength is used.

VI. System Outputs

The system outputs are also defined by membership functions similar to the inputs. The sets for the present pattern are given in Figure 7.

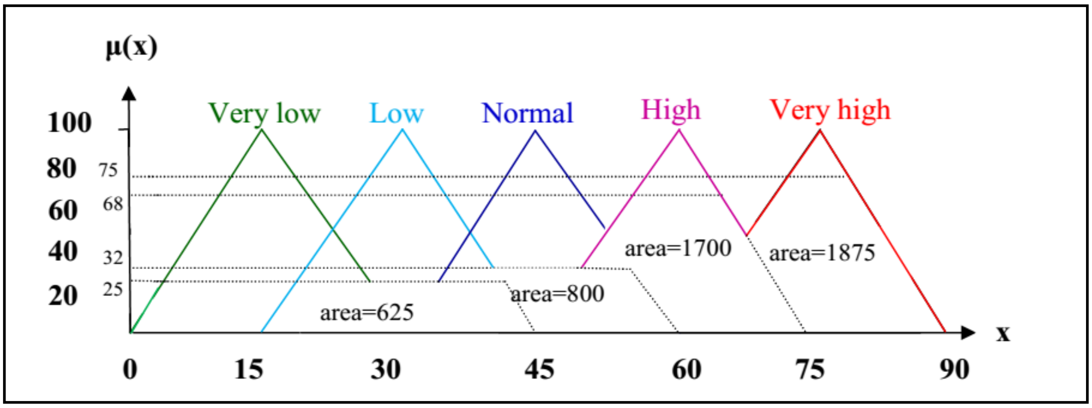

Figure 7: Graph of output membership function of the probability of fog formation

The output membership functions assist in determining a final value of the system output. The procedure is basically the inverse of the evaluation of the degrees of membership. The methodology in use is referred to as the centre-of-gravity method of trapezoidal diagram in which the area of all appropriate output sets, limited in height by the useful rule strength, are evaluated and a weighted mean value of the output is calculated.

VII. Results

In this section will demonstrate an explicit case of the technique as well as indicate some preliminary results using real data. It is indicate the procedure as applied to the current forecast of the probability of fog. The membership assignment above: the dew point, dew point spread, rate of change of the dew point spread and wind speed each involve two fuzzy sets while the sky coverage only involves one set $(2\times2\times2\times2\times1)=1$. There are only 16 rules of the entire set of 144 will have non-zero values, or strengths. The antecedents of all applicable rules along with the predefined corresponding output sets are:

1. IF (moist AND unsaturated AND drying AND strong AND clear) THEN (low probability of fog)
2. IF (moist AND unsaturated AND drying AND excellent AND clear) THEN (medium probability of fog)
3. IF (moist AND unsaturated AND saturating AND strong AND clear) THEN (medium probability of fog)
4. IF (moist AND unsaturated AND saturating AND excellent AND clear) THEN (high probability of fog)
5. IF (moist AND saturated AND drying AND strong AND clear) THEN (medium probability of fog)
6. IF (moist AND saturated AND drying AND excellent AND clear) THEN (very high probability of fog)
7. IF (moist AND saturated AND saturating AND strong AND clear) THEN (very high probability of fog)
8. IF (very moist AND saturated AND saturating AND excellent AND clear) THEN (very high probability of fog)
9. IF (very moist AND unsaturated AND drying AND strong AND clear) THEN (low probability of fog)

Notes

Notes

10. IF (very moist AND unsaturated AND drying AND excellent AND clear) THEN (low probability of fog)
11. IF (very moist AND unsaturated AND saturating AND strong AND clear) THEN (low probability of fog)
12. IF (very moist AND unsaturated AND saturating AND excellent AND clear) THEN (medium probability of fog)
13. IF (very moist AND saturated AND drying AND strong AND clear) THEN (low probability of fog)
14. IF (very moist AND saturated AND drying AND excellent AND clear) THEN (high probability of fog)
15. IF (very moist AND saturated AND saturating AND strong AND clear) THEN (very high probability of fog)
16. IF (very moist AND saturated AND saturating AND excellent AND clear) THEN (very high probability of fog)

Determine the rule strengths are the value of its weakest or least valued antecedent. In the first applicable rule above the value of its antecedents are, in order: 68% (moist dew point), 25% (saturated spread), 28% (drying rate), 28% (excellent wind speed) and 100% (clear sky). The rule strength is then 25. In a similar way, the other rule strengths are evaluated and are listed below

Table 2: Rule Strengths in the current value

Rule	Value	Rule	Value
(i)	25	(ix)	25
(ii)	25	(x)	25
(iii)	25	(xi)	25
(iv)	25	(xii)	25
(v)	28	(xiii)	28
(vi)	28	(xiv)	28
(vii)	28	(xv)	28
(viii)	68	(xvi)	32

This is the applicable rules of the system output for probability of fog formation. The several sets of rules above share the same consequence, or system output. The value of the output is assigned the value of the "most true", or strongest, rule. To illustrate, rules one and two above have as their consequence a low probability of fog formation. Table 2 shows that rules one and two both have a strength of 25, and consequently then the limiting value associated with a low probability of fog is assigned 25. It will be seen next how this limiting value is working. Table 3 shows the values for each of the system output fuzzy sets.

Table 3: Limiting values for each of the system output fuzzy sets

Very Low	Low	Medium	High	Very High
0	25	28	32	68

Figure 7 denotes how these limiting values for each output fuzzy set are used. An individual output membership function is restricted in height by its corresponding limiting value. Then the remaining area of the membership function is calculated. The final system output is then calculated as the weighted average of the centroid of each membership function, with the area of the enclosed set used as the weighting factor. Table 4 shows centroids and weighting areas of the current values.

Table 4: Centroids and weighting areas of output fuzzy sets

Output fuzzy set	very low	low	medium	high	very high
Centroid (C_i)	15	30	45	60	75
Area (A_i)	0	625	800	1700	1875

Hence the final system output $= \dfrac{\sum C_i A_i}{\sum A_i} = 59.5$ which value represents the probability of fog. Finally in this study the fog forecasting is 60%.

Notes

XVII. Conclusion

In this study obtained five input variables are dew point, dew point spread, the rate of change of the dew point spread, wind speed and sky condition at a particular time and output variable is the probability of fog. These five input variables are selected because which is the main factors of the influence the occurrence of fog. Within the context of the current model of fog forecasting, general terms associated with fuzzy logic introduced. The main work that it performs is decision-making. It makes use of the rules of IF and THEN along with the AND connectors which in turn lets it draw some essential rules for decision making. It is a framework which depicts the actual process of converting an input into an output using fuzzy logic. The fuzzy sets are quantitatively defined by membership functions. These functions are typically very simple functions that cover a specified domain of the value of the system input. The functions are generally trapezoids, although simpler functions such as triangles and rectangles and even delta functions are often used. The graphs of the membership function are shown in Figures 2 to 6. The output membership functions assist in determining a final value of the system output. The methodology employed is referred to as the centre-of-gravity method of trapezoidal diagram in which the area of all applicable output sets, limited in height by the applied rule strength, are evaluated and a weighted mean value of the output is calculated. There are only 16 rules of the entire set of 144 will have non-zero values, or strengths. The technique as well as indicate some preliminary results using real data. It is indicate the procedure as applied to the current forecast of the probability of fog. In figure 6, the limiting values for each output fuzzy set are used. An individual output membership function is restricted in height by its corresponding limiting value. Then the remaining area of the membership function is calculated. The final system output is then calculated as the weighted average of the centroid of each membership function, with the area of the enclosed set used as the weighting factor (Table 4). These centroids and weighting areas of the current values are calculated to find the probability of fog. Finally, this study attempted to forecast the fog based on Situation of fog parameters and fuzzy logic.

References Références Referencias

1. T. S. Glickman, Glossary of Meteorology, 2nd edition, *American Meteorological society*, 2000: 855.
2. I. Gultepe, R. Tardif, S. C. Michaelides, J. Cermak, A. Bott, J. Bendix, M. D. Muller, M. Pagowski, B. Hansen, G. Ellrod, W. Jacobs, G. Toth and S. G. Cober,

Notes

Fog research: A review of past achievements and future perspectives, *Pure and Applied Geophysics,* 2007, 164:1121–1159.

3. H. Tanaka, S. Honma, M. Nishi, T. Igarashi, S. Teramoto, F. Nishio and S. Abe, Acid fog and hospital visits for asthma: an epidemiological study, *European Respiratory Journal*, 1998, 11(6):1301–1306.
4. D. Liu, Z. Li, W. Yan, and Y. Li, Advances in fog microphysics research in China, *Asia Pacific Journal of Atmospheric Sciences,* 2017, 53(1):1–18.
5. Z. Li, J. Yang, C. Shi and M. Pu, Urbanization effects on fog in china: field research and modeling, *Pure and Applied Geophysics,* 2012, 169(5):927–939.
6. T. W. Choularton, G. Fullarton, J. Latham, C. S. Mill, M. H. Smith and I. M. Stromberg A field study of radiation fog in Meppen, West Germany", Quarterly Journal of the Royal Meteorological Society, (1981), "107(452), 381–394.
7. M. Mohan and S. Payra, Influence of aerosol spectrum and air pollutants on fog formation in urban environment of megacity Delhi, India, *Environmental Monitoring and Assessment,* 2009, 151(1-4):265–277.
8. H. Kohler, The nucleus in and the growth of hygroscopic droplets, *Transactions of the Faraday Society,* 1936, 32:1152–1161.
9. B. Zhou Introduction to a new fog diagnostic scheme, *NCEP Office Note,* 2011, 466, 43. Online at: www.emc.ncep.noaa.gov/officenotes/newernotes/on466.pdf .
10. H. Hu, Q. Zhang, B. Xie, Y. Ying, J. Zhang and X. Wang, Predictability of an advection fog event over north China. Part I: sensitivity to initial condition differences, *Mon. Wea. Rev.,* (2014), 142(5):1803–1822.
11. M. Jim, Application of fuzzy logic in operational meteorology, *Carlifornia: Addison Wesley*, Longman Inc., 2005.
12. G. Viot, Fuzzy Logic in C, *Dr. Dobb's Journal*, special issue on Cognitive Computing, 1993, 40-49.
13. L. Christian, C. Mwongera, P. Camberlin and J. B. Micheau, Indigenous Past Climate Knowledge as Cultural Built-in Object and Its Accuracy, *Ecology and Society,* 2013, 18(4):22-45.
14. T. J. Ross, Fuzzy Logic with Engineering Applications, *McGraw-Hill International Editions,* 1997.
15. A. H. Agboola, A. J. Gabriel, E. O. Aliyu and B. K. Alese, Development of a fuzzy logic based rainfall prediction model, *International Journal of Engineering and Technology*, 2013, 3(4):427-435.
16. M. A. Rahman, Improvement of Rainfall Prediction Model by Using Fuzzy Logic, *American Journal of Climate Change*, 2020, 9(4):391-399.
17. A. Hasan and M. A. Rahman, "A Study on Rainfall Calibration and Estimation at the Northern Part of Bangladesh by Using Mamdani Fuzzy Inference System, *Journal of Environment Protection and Sustainable Development,* 2019, 5(2):58-69.
18. S. Sujitjorn, P. Sookjaras and W. Wainikorn, An expert system to forecast visibility in Don Muang Air Force Base", IEEE Internat, Conf. on Systems, Man and Cybernetics Humans, *Information and Technology*, IEEE, NY, USA, 1994, 2528–2531.
19. J. Murtha, Applications of fuzzy logic in operational meteorology, *Scientific Services and Professional Development Newsletter*, Canadian Forces Weather Service, 1995, 42–54.

20. K. Petty, B. Carmichel, G. Wiener, M. Petty and M. A. Limber, Fuzzy logic system for the analysis and prediction of cloud ceiling and visibility, Preprints Ninth Conference on Aviation, Range and Aerospace Meteorology, Orlando, Fl., *American Meteorological Society*, 2000, 331–333.
21. B. K. Hansen, Analog forecasting of ceiling and visibility using fuzzy sets, 2nd Conference on Artificial Intelligence, *American Meteorological Society,* 2000, 1–7.
22. L. A. Zadeh, Fuzzy Sets, Information and Control, 1965, 8:338- 353.
23. J. Li, J. Sun, and M. Zhou, Observational analyses of dramatic developments of a severe air pollution event in the Beijing area, *Atmospheric Chemistry and Physics,* 2018, 18(6):3919–3935.

Notes

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 21 Issue 5 Version 1.0 Year 2021
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

Estimating the Proportion of True Null Hypotheses: A Likelihood Approach

By Hualing Zhao & Hanfeng Chen

Wuhan University of Technology

Abstract- Many estimators for the proportion π_0 of the true null hypotheses in a multiple testing problem have been proposed in literature. Motivated from the work on the histogram approach, in this article we propose a new estimator based on the likelihood function with an approximating alternative histogram. AIC is used to select the number of bins for the histogram. Simulation study demonstrates that the new estimator outperforms and substantially improves existing methods including Storey estimators, convex density estimator, and histogram estimator. The new method is applied to a real-life data set of breast cancer.

Keywords: *akaike information criterion; false discovery rate; finite mixture model; multiple comparisons.*

GJSFR-F Classification: *MSC 2010: 97K80*

ESTIMATINGTHEPROPORTIONOFTRUENULLHYPOTHESESALIKELIHOODAPPROACH

Strictly as per the compliance and regulations of:

Notes

Estimating the Proportion of True Null Hypotheses: A Likelihood Approach

Hualing Zhao [α] & Hanfeng Chen [σ]

Abstract- Many estimators for the proportion π_0 of the true null hypotheses in a multiple testing problem have been proposed in literature. Motivated from the work on the histogram approach, in this article we propose a new estimator based on the likelihood function with an approximating alternative histogram. AIC is used to select the number of bins for the histogram. Simulation study demonstrates that the new estimator outperforms and substantially improves existing methods including Storey estimators, convex density estimator, and histogram estimator. The new method is applied to a real-life data set of breast cancer.

Keywords: akaike information criterion; false discovery rate; finite mixture model; multiple comparisons.

I. Introduction

Consider the problem of estimating the proportion π_0 of true null hypotheses in a collection of m tests, given the observed p-values $p_1, \cdots, p_m$ for the collection of m tests. This problem has attracted a lot of attentions in statistical literatures, attributing to its important role in dealing with multiple testing procedures, since the nominal paper Storey (2002). It has naturally arisen in assessing or controlling an overall false rejection rate, i.e., the false discovery rate (FDR) proposed by Benjamini and Hochberg (1995) in multiple null hypothesis testing problems and multiple comparisons. Benjamini and Hochberg proved that if $p_1, \cdots, p_m$ are independent with continuous distributions, the popular Sime 's multiple testing procedure (Sime 1986), where a null hypothesis is rejected whenever the observed p-value is less than α, results in an FDR controlled by $\pi_0\alpha$. Therefore, a reliable estimate is essential to control the FDR in the Sime's multiple testing procedure and improve its testing power. On the other hand, the proportion π_0 is a quantity of interest in its own right. For example, researchers may wish to estimate the proportion of genes that are not differentially expressed in DNA microarray experiments (Parker et. al. 1988).

Store (2002) proposed to approach the estimating problem from a Bayesian point of view by treating $p_1, \cdots, p_m$ as a random sample of size m from a mixture distribution of a uniform distribution and a non-uniform distribution with mixing proportion π_0. Specifically, consider a population consisting of all outcomes of the p-values $p_1, \cdots, p_m$ in testing m null hypotheses. There are two types of p-values: true-null p-values and false-null p-values. Following Tong, Feng, Hilton and Zhao (2013)'s terminology, by true-null p-value we mean the p-value is observed from a true null hypothesis and by false-null p-value we mean it is observed from a false null hypothesis. Thus, the population distribution can be described as a finite mixture with mixing proportion π_0 of the uni-

Author α: Department of Statistics, School of Sciences, Wuhan University of Technology, Wuhan, Hubei 430070, China.
e-mail: zhaohualing2011@whut.edu.cn
Author σ: Department of Mathematics and Statistics, Bowling Green State University, Bowling, Green, OH 43403, USA.
e-mail: hchen@bgsu.edu

form distribution on $[0, 1]$ and another non-uniform distribution with the pdf, say $h(x)$, with respect to the Lebesque measure dx. Denote the mixture by $f(x|\pi_0, h)$, i.e.,

$$f(x|\pi_0, h) = \pi_0 + (1 - \pi_0)h(x), \quad 0 \leq x \leq 1. \tag{1}$$

By Storey's approach, the observed p-values are considered to be a random sample of size m from $f(x|\pi_0, h)$. As a consequence, the mixing proportion π_0 in the mixture model (1) represents the proportion of true-null p-value subpopulation and an estimate for π_0 based on a random sample from (1) thus yields an estimate for the proportion of true null hypotheses among the m null hypotheses in the multiple testing problem.

Notes

The first well-studied estimator for π_0 was proposed by Storey (2002) as follows:

$$\hat{\pi}_0^s(\lambda) = W(\lambda)/\{m(1 - \lambda)\},$$

where λ is an appropriately chosen value close to 1 and $W(\lambda)$ is the total number of the p-values greater than λ. This estimator is motivated by the fact that false-null p-values are typically small so that with λ close to 1, $1 - \lambda$ is close to the expected proportion of true-null p-values falling into $(\lambda, 1]$. Since the proportion of the p-values falling into $(\lambda, 1]$ is a mixture of $1 - \lambda$ and $\int_\lambda^1 h(x)dx$ with mixing proportions π_0, $\hat{\pi}_0^s$ is obtained mathematically by simply setting the proportion of false-null p-values falling into $(\lambda, 1]$ to be zero, i.e., $\int_\lambda^1 h(x)dx = 0$. As a consequence, $\hat{\pi}_0^s$ tends to overestimate π_0, as $\int_\lambda^1 h(x)dx$ is typically positive. The biasedness can be too great to be acceptable when π_0 belongs to the lower part of $[0, 1]$. Storey (2002) proposed a bootstrap procedure to choose λ that minimizes an upper bound of the mean square error of the resulting positive FDR estimator $\hat{Q}(\lambda) = \alpha\hat{\pi}_0^s(\lambda)/R(\alpha)$ where $R(\alpha) = \#\{p_i \leq \alpha\}$. See more discussion and some other work on how to select λ, see Storey, Taylor and Siegmund (2004), and Nettleton, Hwang, Caldo and Wise (2006).

Since the publication of Storey (2002), the mixture model approach described above has been adopted widely and many other estimators have been proposed. They include: Langaas, Lindqvist and Ferkingstad (2005)'s by a histogram approach, Wu, Guan and Zhao (2006)'s polynomial-type estimator, Jiang and Doerge (2008)'s estimator averaging over different λ-values with Storey (2002)'s estimator, Zhao, Wu, Zhang and Chen (2012)'s estimator in exponential mixture model, Cheng, Gao and Tong (2015)'s estimator with reduction of estimating bias and standard deviation. For more references and discussion, see Cheng, Gao and Tong (2015), and Tong, Feng, Hilton and Zhao (2013) that deals with possibly dependent p-values. Nevertheless, when these estimators are intended to improve some aspects of Storey's estimator, the biasedness remains significant even with m as large as 2,000, especially when π_0 is not close to 1. Motivated by the histogram approach (see Mosig et al. 2001 and Nettleton et al. 2006), a new estimator is proposed in this paper via a likelihood approach with h being approximated by a modified histogram pdf. Akaike information criterion is used to select the number of categories in histogram construction. Simulation study demonstrates that the new estimator significantly improves popularly cited existing methods such as Storey (2002), Langaas, Lindqvist and Ferkingstad Langaas (2005), Jiang and Doerge (2008) and Nettleton et al.(2006).

The paper is organized as follows. The new method along with discussion of computing issues is presented in Section 2. Simulation results are reported in Section 3. A real-life example is given in Section 4.

II. Method

Let $p_1, ..., p_m$ be a random sample of size m from the pdf $f(x|\pi_0, h)$ defined in (1). We propose π_0 to be estimated by the MLE when h is subject to a histogram-type approximation. We shall proceed with a study on the identification problem in the mixture model.

Notes

a) *Identification*

Let Θ be the parameter space consisting of all (π_0, h)'s with $\pi_0 \in (0, 1)$ and h being a probability density function on [0,1] and continuous in a neighborhood of 1with $h(1-) = 0$, where

$$h(1-) = \lim_{x \to 1-} h(x).$$

For convenience and confirmation, when a pdf is used as a parameter throughout the paper, we should refer to the probability distribution, say the CDF specified by the pdf.

Lemma 1 The distribution $f(x|\pi_0, h)$ in the model (1) is identifiable in $(\pi_0, h) \in \Theta$.

Proof. Let two parameter sets (π_1, h_1) and (π_2, h_2) in Θ define the same distribution, i.e., for any $x \in [0, 1]$,

$$\pi_1 x + (1 - \pi_1) H_1(x) = \pi_2 x + (1 - \pi_2) H_2(x), \tag{2}$$

where H_i is the CDF of h_i, $i = 1, 2$. Since h_1 and h_2 are continuous in a common neighborhood of 1, (2) implies that for any x in a neighborhood of 1,

$$\pi_1 + (1 - \pi_1) h_1(x) = \pi_2 + (1 - \pi_2) h_2(x).$$

So $\pi_1 = \pi_2$ because $h_1(1-) = h_2(1-) = 0$. This in turn implies $H_1 = H_2$ in (2) because $\pi_1 < 1$ and $\pi_2 < 1$. The lemma is proved.

Remarks.

(a) First of all, note the simple fact that $h(1-) = 0$ implies that h is non-uniform. Of course, identification of the parameter π_0 of interest requires that h is non-uniform.

(b) When $\pi_0 = 1$, f is un-identifiable in h.

(c) The assumption that h is continuous in a neighborhood of 1 with $h(1-) = 0$ is critical. Without it, the lemma can fail. For a counterexample, let $\pi_1 = 1/3$ and $h_1(x) = 1/4 + (3/2)x$, and $\pi_2 = 1/2$ and $h_2(x) = 2x$. We have

$$\int_0^1 h_1(x)dx = 1/4 + 3/4 = 1, \int_0^1 h_2(x)dx = 1,$$

and for any x in $[0, 1]$

$$\pi_1 + (1 - \pi_1) h_1(x) = 1/3 + (2/3)[1/4 + (3/2)x] = 1/2 + x,$$

and

$$\pi_2 + (1-\pi_2)h_2(x) = 1/2 + (1/2)(2x) = 1/2 + x.$$

That is, $\pi_1 + (1-\pi_1)h_1(x) = \pi_2 + (1-\pi_2)h_2(x)$ for all x in $[0,1]$, but $\pi_1 \neq \pi_2$ and $h_1 \neq h_2$.

b) A histogram approximation to h

Motivated by the histogram approach (see Mosig et al. (2001), a histogram approximation to the alternative pdf $h(x)$ is proposed as follows. Let $k > 2$ be an integer. Define

$$\tilde{h}(x) = \begin{cases} kq_j, & \text{if } (j-1)/k \leq x < j/k,\ 1 \leq j \leq k-1 \\ k^2 q_{k-1}(1-x), & \text{if} (k-1)/k \leq x \leq 1 \end{cases}$$

where $0 \leq q_1 \leq 1, \cdots, 0 \leq q_{k-1} \leq 1$ satisfying $q_1 + \ldots + q_{k-1} + q_{k-1}/2 = 1$. The linear modification over the right most subinterval $(1-1/k, 1]$ warrants $\hat{h} \in \Theta$ so that $f(x|\pi_0, \tilde{h})$ is identifiable in π_0 and $q = (q_1, \cdots, q_{k-2})$. As Storey(2002) and Tong et al. (2013) remarked, the false-null p-values that follow the distribution h are typically small. In other words it is often the case in application that the alternative pdf h is highly skewed to the right. Therefore, with the linear modification on the rightmost subinterval $(1-1/k, 1]$, the effect on the accuracy of the estimate for Storey (2002), Jiang and Doerge (2008) and some other methods assume that h is zero in a neighborhood of 1.

c) The new estimator with fixed k

Suppose that $k > 2$ is specific for now (a selection procedure will be described later). When h is approximated by or restricted to $\tilde{h}$, with a random sample $p_1, \cdots, p_m$, the log-likelihood of the parameter π_0 of interest and the new nuisance parameter q becomes

$$\begin{aligned} l(\pi_0, q) &= \sum_{i=1}^{m} \log f(p_i|\pi_0, \tilde{h}) \\ &= \sum_{i=1}^{m} \log \left\{ \pi_0 + (1-\pi_0)\{\prod_{j=1}^{k-1} (kq_j)^{\xi_{ij}}\}\{k^2 q_{k-1}(1-p_i)\}^{\xi_{ik}}\} \right\}, \end{aligned} \tag{3}$$

where ξ_{ij} is the indicator whether p_i falls into the j-th category of the histogram with k bins or not, i.e.,

$$\xi_{ij} = \begin{matrix} 1, & \text{if } (j-1)/k \leq p_i < j/k, \\ 0, & \text{otherwise} \end{matrix} \tag{4}$$

for $i = 1, \cdots, m$, $j = 1, \cdots, k$.

Let $(\hat{\pi}_0(k), \hat{q})$ be the MLE of (π_0, q) with the log-likelihood function (3), subject to $0 < \pi_0 < 1$ and $0 < q_1 < 1, \cdots, 0 < q_{k-1} < 1$, $q_k = q_{k-1}/2$ such that $\sum_{j=1}^{k} q_j = 1$. This defines the new estimator $\hat{\pi}_0(k)$ for π_0 with specified k. A direct application of the standard theory of MLE yields the following theorem and hence a theory of consistency of the new estimator.

Notes

Theorem 1 *For any* $(\pi_0, h) \in \Theta$ *and* $k > 2$, *let*

$$q_j = \int_{(j-1)/k}^{j/k} h(x)dx$$

Notes

for $j = 1, \cdots, k-2$. *Put* $q = (q_1, \cdots, q_{k-2})$ *and* $\theta = (\pi_0, q)$. *Then* $\sqrt{m}(\hat{\pi}_0(k) - \pi_0, \hat{q} - q)$ *converges to* $N(0, \Sigma)$ *in distribution, as* $m \to \infty$, *where*

$$\Sigma = -\left[E\left(\frac{\partial^2}{\partial\theta\partial\theta}\log f(X|\pi_0, \tilde{h})\right)\right]^{-1}.$$

d) *Computation of* $\hat{\pi}_0(k)$

Maximizing the nonlinear log-likelihood function (3) can be complicating. However, the EM algorithm can be used to obtain an approximation to the MLE $\hat{\pi}_0(k)$ easily. To do so, introduce a latent Bernoulli variable w indicating the component-ship of the p-value in the finite mixture distribution. That is, let w be a binary random variable with $P(w = 1) = \pi_0$ and $P(w = 0) = 1 - \pi_0$, and let a random variable p follow the distribution as follows. Given $w = 1$, p follows the uniform distribution on $[0, 1]$ and given $w = 0$, $p \sim \tilde{h}$. It is clear that p follows the mixture distribution $f(x|\pi_0, \tilde{h})$. Thus $\mathcal{P} = \{p_1, \cdots, p_m\}$ can be viewed as the incomplete data of a random sample $(p_1, w_1), \cdots, (p_m, w_m)$ from (p, w) with missing values $w_1, \cdots, w_m$.

Note that with the complete data $(p_1, w_1), \cdots, (p_m, w_m)$, the likelihood function of (π_0, q) is

$$\prod_{i=1}^{m} \pi_0^{w_i}\left\{(1-\pi_0)[\prod_{j=1}^{k-1}(kq_j)^{\xi_{ij}}][k^2 q_{k-1}(1-p_i)]^{\xi_{ik}}\right\}^{1-w_i}, \tag{5}$$

where ξ_{ij}'s are defined in (4), and the log-likelihood function is thus

$$l^*(\pi_0, q) = w_{\cdot}\log\pi_0 + (m - w_{\cdot})\log(1-\pi_0) + \sum_{j=1}^{k-1}\xi^*_{\cdot j}\log(kq_j) + \sum_{i=1}^{m}\xi^*_{ik}\log[k^2 q_{k-1}(1-p_i)], \tag{6}$$

where $\xi^*_{ij} = (1 - w_i)\xi_{ij}$, $w_{\cdot} = \sum_{i=1}^m w_i$, and $\xi^*_{\cdot j} = \sum_{i=1}^m \xi^*_{ij}$.

The EM algorithm can be easily implemented as follows. Let $(\pi_0^{(s)},\ q^{(s)})$ be the current approximations to the MLE $(\hat{\pi}_0(k), \hat{q})$ with the log-likelihood function l given in (3). The next approximation $(\pi_0^{(s+1)},\ q^{(s+1)})$ is given by the EM algorithm in two steps, the so-called E-step and M-step.

E-Step: Compute the conditional expectation of the log-likelihood function

$$Q(\pi_0, q) = E_{\pi_0^{(s)}, q^{(s)}}\{l^*(\pi_0, q)|\mathcal{P}\}$$

$$= E_{\pi_0^{(s)},q^{(s)}}(w_.|\mathcal{P})\log(\pi_0) + (m - E_{\pi_0^{(s)},q^{(s)}}(w_.|\mathcal{P}))\log(1-\pi_0)$$

$$+\textstyle\sum_{j=1}^{k-1} E_{\pi_0^{(s)},q^{(s)}}(\xi^*_{.j}|\mathcal{P}))\log(kq_j) + \sum_{i=1}^{m}\{E\{\xi^*_{ik}\log[k^2 q_{k-1}(1-p_i)]|\mathcal{P}\}.$$

Note

$$\begin{aligned} E_{\pi_0^{(s)},q^{(s)}}(w_i|\mathcal{P}) := \hat{w}_i &= P_{\pi_0^{(s)},q^{(s)}}(w_i = 1|\mathcal{P}) \\ &= \frac{\pi_0^{(s)}}{\pi_0^{(s)} + (1-\pi_0^{(s)})\{\prod_{j=1}^{k-1}(kq_j^{(s)})^{\xi_{ij}}\}\{k^2 q_{k-1}^{(s)}(1-p_i)\}^{\xi_{ik}}}. \end{aligned}$$

We have

$$\begin{aligned} Q(\pi_0, q) &= (\sum_{i=1}^{m}\hat{w}_i)\log\pi_0) + (m - \sum_{i=1}^{m}\hat{w}_i)\log(1-\pi_0) + \sum_{j=1}^{k-1}\left[\sum_{i=1}^{m}(1-\hat{w}_i)\xi_{ij}\right]\log(q_j k) \\ &+ \sum_{i=1}^{m}(1-\hat{w}_i)\xi_{ik}\log[k^2 q_{k-1}(1-p_i)]. \end{aligned}$$

M-Step: In the M-step, $Q(\pi_0, q)$ is maximized to yield the next approximation $\pi_0^{(s+1)}$ and $q^{(s+1)}$. Setting $\partial Q/\partial\pi_0 = 0$, we immediately have

$$\pi_0^{(s+1)} = \frac{\sum_{i=1}^{m}\hat{w}_i}{m}.$$

By Lemma 2 in Appendix,

$$\begin{aligned} q_j^{(s+1)} &= \frac{1}{\hat{m}}\sum_{i=1}^{m}(1-\hat{w}_i)\xi_{ij},\ j = 1,\cdots,k-2 \\ q_{k-1}^{(s+1)} &= \frac{2}{3\hat{m}}\sum_{i=1}^{m}(1-\hat{w}_i)(\xi_{i(k-1)} + \xi_{ik}) \\ q_k^{(s+1)} &= q_{k-1}^{(s+1)}/2, \end{aligned}$$

The formulas in the EM algorithm above are similar to to those developed by Oluyemi (2016) and Oluyemi and Chen (2016) where there are some algebraic errors.

where

$$\hat{m} = \sum_{j=1}^{k}\sum_{i=1}^{m}(1-\hat{w}_i)\xi_{ij} = m - \sum_{i=1}^{m}\hat{w}_i\ .$$

Notes

Notes

e) Selection of k via AIC

In this subsection, we discuss about selection of k. As a histogram-type approximation to $h(x)$, the value of k can reveal different features of the data. Since k is the number of categories of the histogram fitting h, a larger value of k does a better fitting job and hence is expected to result in a more accurate estimate of π_0. However, a larger value of k causes a greater standard error in estimating π_0 because there are more nuisance parameters to handle. As the estimate $\hat{\pi}_0(k)$ is based on the likelihood function, the Akaike information criterion (AIC) can be a nature choice of criteria for selection of an appropriate value of k. Let $\hat{l}_k$ be the maximum value of the log-likelihood function l defined in (3), i.e., $\hat{l}_k = l(\hat{\pi}_0(k), \hat{q})$. Noting that we have total of $k-1$ free parameters in computing $\hat{l}_k$, the AIC selection of k is to choose $\hat{k}$ such that $2\hat{l}_k - 2(k-1)$ is maximized, i.e.,

$$\hat{k} = \arg\max\{2\hat{l}_k - 2(k-1)\}.$$

The final estimate $\hat{\pi}_0$ for π_0 is $\hat{\pi}_0 = \hat{\pi}_0(\hat{k})$.

III. Simulation Study

Simulation studies are conducted with the p-values based on one-sided z-test in finite normal mixture models to evaluate the performance of the new estimator $\hat{\pi}_0$ and compare it with some existing method. Specifically, consider the finite normal model $\pi_0 N(0,1) + (1-\pi_0)N(1,1)$ with five griding values of π: 0.2, 0.35, 0.50, 0.65 and 0.8 . The p-value is computed by $p = 1 - \Phi(z)$. Four sample sizes, $m = 500, 1000, 1500$ and 2000 are consider. With each combination, 2000 Monte Care trials are used.

In the simulation studies, the EM algorithm is used to approximate the MLE's and a linear algorithm is used to search for the AIC selection $\hat{k}$.

Four existing estimators popularly cited in the literature are considered for purpose of comparison. They are Storey's estimator with bootstrap method $\hat{\pi}_0^s$, the convex density estimator $\hat{\pi}_0^c$ by Langaas, Lindqvist and Ferkingstad (2005), the averaging Storey estimator $\hat{\pi}_0^a$ by Jiang and Doerge (2008) and the histogram estimator $\hat{\pi}_0^h$ by Nettleton, Hwang, Caldo, and Wise (2006). The R package **cp4p** is used for computations of these estimates; all the computations are done with default settings of the R package.

Simulation results are reported in Table 1. The following findings from the simulation studies are immediate:

1. The new estimator $\hat{\pi}_0$ performs very well. A clear converging pattern is demonstrated in each simulation model, as m increases. The convergence with lower values of π_0 is faster than with higher values.
2. The new estimator performs substantially better than all other four existing estimators. It is noted that the Storey's $\hat{\pi}_0^s$ performs somewhat better than the other three existing estimators in general and its best performance is on the higher values of π_0 as expected. See Storey (2002).

IV. Real Data Analysis

In this section, we apply the new estimate method to the real life data from Hedenfalk et al. (2001)where 3226 genes were studied with $n_1 = 7$ BRECA1 arrays and $n_2 = 8$ BRCA2 arrays. The example is to test with each gene the null hypothesis that there is no differential gene expression between BRCA1-mutation-positive tumors and BRCA2-mutation-positive tumors by using a

Table 1: Empirical average of the estimates for the proportion π_0 with their empirical standard deviations in parentheses. The true model for the p-value of one-sided test is: $p = 1 - \Phi(x)$ with $x \sim \pi_0 N(0,1) + (1-\pi_0)N(1,1)$. Each of the entries is based on 2,000 Monte Carlo trials.

m	π_0	$\hat{\pi}_0$	$\hat{\pi}_0^s$	$\hat{\pi}_0^c$	$\hat{\pi}_0^a$	$\hat{\pi}_0^h$
500	0.20	0.194	0.339	0.324	0.357	0.378
		(0.101)	(0.059)	(0.061)	(0.068)	(0.593)
500	0.35	0.327	0.454	0.569	0.488	0.507
		(0.116)	(0.066)	(0.066)	(0.067)	(0.096)
500	0.50	0.457	0.586	0.583	0.618	0.644
		(0.130)	(0.067)	(0.065)	(0.066)	(0.086)
500	0.65	0.605	0.710	0.712	0.746	0.773
		(0.123)	(0.069)	(0.064)	(0.060)	(0.068)
500	0.80	0.749	0.826	0.834	0.864	0.885
		(0.113)	(0.066)	(0.059)	(0.052)	(0.051)
1000	0.20	0.202	0.328	0.308	0.322	0.314
		(0.082)	(0.044)	(0.047)	(0.057)	(0.085)
1000	0.35	0.336	0.455	0.445	0.462	0.468
		(0.095)	(0.049)	(0.051)	(0.060)	(0.086)
1000	0.50	0.480	0.583	0.578	0.599	0.615
		(0.093)	(0.052)	(0.050)	(0.055)	(0.074)
1000	0.65	0.624	0.701	0.707	0.729	0.750
		(0.089)	(0.052)	(0.049)	(0.052)	(0.065)
1000	0.80	0.777	0.829	0.834	0.886	0.875
		(0.084)	(0.050)	(0.044)	(0.041)	(0.042)
1500	0.20	0.206	0.292	0.291	0.298	0.294
		(0.072)	(0.042)	(0.040)	(0.051)	(0.071)
1500	0.35	0.348	0.423	0.433	0.440	0.449
		(0.079)	(0.049)	(0.044)	(0.056)	(0.074)
1500	0.50	0.480	0.553	0.569	0.581	0.579
		(0.081)	(0.053)	(0.043)	(0.054)	(0.070)
1500	0.65	0.636	0.684	0.703	0.721	0.741
		(0.075)	(0.055)	(0.043)	(0.048)	(0.055)
1500	0.80	0.788	0.801	0.829	0.848	0.870
		(0.068)	(0.050)	(0.040)	(0.039)	(0.042)
2000	0.20	0.206	0.295	0.290	0.292	0.284
		(0.066)	(0.038)	(0.035)	(0.045)	(0.062)
2000	0.35	0.348	0.425	0.434	0.435	0.436
		(0.071)	(0.042)	(0.037)	(0.050)	(0.067)
2000	0.50	0.496	0.556	0.571	0.579	0.585
		(0.065)	(0.047)	(0.038)	(0.051)	(0.065)
2000	0.65	0.636	0.683	0.700	0.713	0.731
		(0.071)	(0.050)	(0.038)	(0.047)	(0.056)
2000	0.80	0.798	0.812	0.831	0.847	0.865
		(0.056)	(0.047)	(0.035)	(0.035)	(0.038)

Notes

Notes

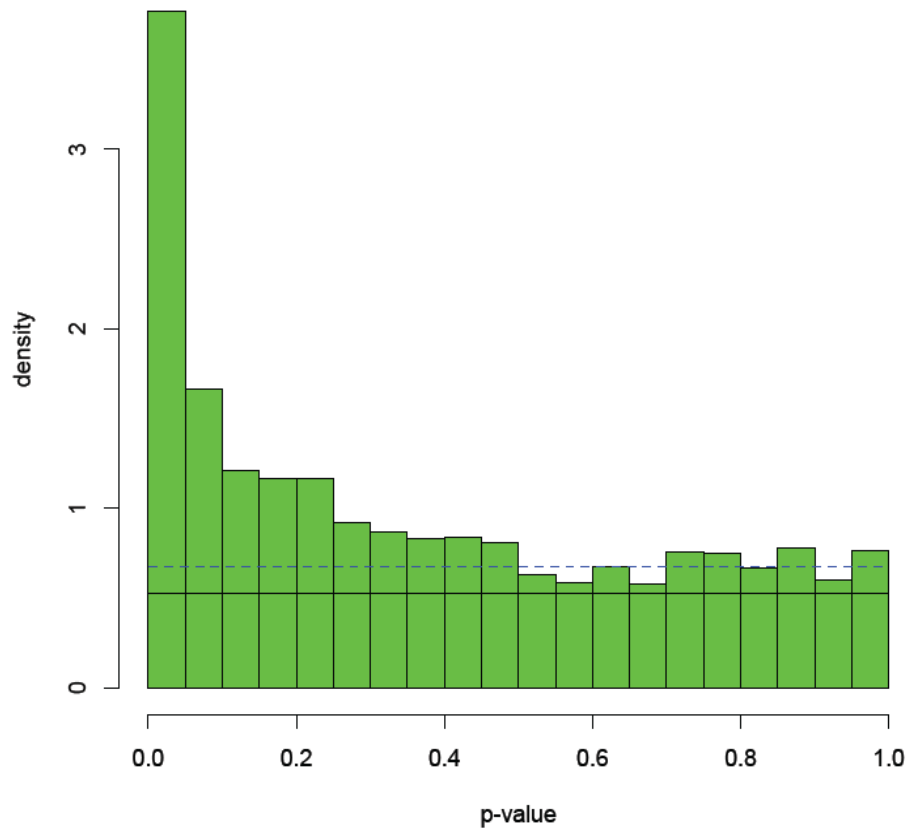

Figure 1: Histogram of the available 2752 p-values for detection of differently expressed genes with the data in Hedenfalk et al.(2001). The solid line has the intercept of $\hat{\pi}_0 = 0.523$ and the dash line has the intercept of $\hat{\pi}_0^h = 0.677$.

two-sample t-statistic. It was also analyzed by Storey and Tibshirani (2003). Following their instructions, we were able to download the p-values from *http://www.genomine.org/qvalue/results.txt.* Removing missing values left with m=2752 p-values.

The estimates for the proportion π_0 of true null hypotheses by the new estimator and the other four existing estimators are listed in Table 2. The AIC selection of k for the new estimator is $\hat{k} = 41$. In Figure 1 for the histogram of the 2752 available p-values with data in Hedenfalk et al. (2001), the solid horizontal line would be the expected bottom of the graph of the density function of the p-value under assumption $\inf h(x) = 0$ if the real value of π_0 were equal to the estimate $\hat{\pi}_0$ (0.523), whereas the dash line would indicate the expected bottom of the graph of the density function if the real value of π_0 were equal to the histogram estimate $\hat{\pi}_0^h = 0.677$. It is evident from Figure 1 that the existing estimators overestimates π_0.

Table 2: Estimates for the proportion π_0 of not differentially expressed genes with the break cancer data in Hedenfalk et al. (2001)

$\hat{\pi}_0$	$\hat{\pi}_0^s$	$\hat{\pi}_0^c$	$\hat{\pi}_0^a$	$\hat{\pi}_0^h$
0.523	0.689	0.682	0.704	0.677

References Références Referencias

1. Storey, J.D. (2002) A direct approach to false discovery rates. *J. R. Stat. Soc. Ser. B* 64(3): 479-498.

2. Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. Roy. Stat. Soc. B* 57(1): 289-300.
3. Simes, R.J. (1986) An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 73: 751-754.
4. Parker, R.A. & Rothenberg, RB. (1988) Identifying important results from multiple statistical tests. *Statistics in Medicine* 7: 1031-1043.
5. Tong, T., Feng, Z., Hilton, J.S. & Zhao, H. (2013) Estimating the proportion of true null hypotheses using the pattern of observed p-values. *J. Appl. Stat.* 40(9): 1949-1964.
6. Storey, J.D., Taylo, J.E. & Siegmund, D. (2004) Strong control, conservative point estimation, and simultaneous conservative consistency of false discovery rates. *J. R. Stat. Soc. Ser. B* 66(1):187-205.
7. Nettleton, D., Hwang, J.T.G., Caldo, R.A. & Wise, R.P. (2006) Estimating the number of true null hypotheses from a histogram of p-values. *Journal of Agricultural, Biological, and Environmental Statistics* 11(3): 337-356.
8. Zhao, H., Wu, X., Zhang, H. & Chen, H. (2012) Estimating the proportion of true null hypotheses in nonparametric exponential mixture model with application to the leukemia gene expression data. *Communications in Statistics-Simulation and Computation* 41(9): 1580–1592.
9. Langaas, M. & Lindqvist, B. H. & Ferkingstad, E. (2005) Estimating the proportion of true hypotheses, with application to DNA microarray data. *J. R. Stat. Soc. B* 67(4): 555-572.
10. Wu, B., Guan, Z. & a Zhao, H. (2006) Parametric and nonparametric FDR estimation revisited. *Biometrics* 62(3): 735-744.
11. Cheng, Y., Gao, D. & Tong, T. (2015) Bias and variance reduction in estimating the proportion of true-null hypotheses. *Biostatistics* 16(1): 189-204.
12. Mosig, M. O., Lipkin, E. Khutoreskaya, G. Tchourzyna, E., Soller, M. & and Friedmann, A. (2001) A whole genome scan for quantitative trait loci affecting milk protein percentage in Israeli-Holstein cattle, by means of selective milk DNA pooling in a Daughter design, using an adjusted false discovery rate criterion. *Genetics* 157: 1683-1698.
13. Jiang, H., & Doerge, R. (2008) Estimating the proportion of true null hypotheses for multiple comparisons. *Cancer Informatics* 6: 25-32.
14. Hedenfalk, I., Duggan, D. Chen, Y.D., Radmacher, M., Bittner, M., Simon, R., Meltzer, P., Gusterson, B., Esteller, M., Kallioniemi, O.P., Wilfond, Borg, A., & Trent, I. (2001) Gene expression profiles in hereditary breast cancer. *New England Journal of Medicine* 344(8): 539-548.
15. Oyeniran, O. (2016) Estimating the proportion of true null hypotheses in multiple testing problems. *Doctoral Dissertation, Bowling Green State University.*
16. Oyeniran, O. & Chen, H. (2016) Estimating the proportion of true null hypotheses in multiple testing problems. *Journal of Probability and Statistics.* Volume 2016, Article ID 3937056, http://dx.doi.org/10.1155/2016/3937056.
17. Storey, J.D. (2003) The positive false discovery rate: a Bayesian interpretation and the q-value. *Ann. Stat.* 31(6): 2013-2035.
18. Storey, J.D. & Tibshirani, R. (2003) Statistical significance for genome wide studies. *Proc. Nat. Acad. Sci.s* 100(16): 9440-9445.

Notes

Appendix: Computation for the estimate of q

In M-step of the EM iteration algorithm, we compute the update for the estimate of q based on the following result.

Lemma 2 Let $g(q) = \sum_{j=1}^{k-1} x_j \log q_j$, where $x_j \geq 0$ are constant and $0 \leq q_j \leq 1$ such that

Notes

$$\sum_{j=1}^{k-2} q_j + (3/2)q_{k-1} = 1.$$

Then the maximum of g is attained at $\hat{q}_1 = x_1/\hat{m}, \cdots, \hat{q}_{k-2} = x_{k-2}/\hat{m}$, $\hat{q}_{k-1} = (2/3)x_{k-1}/\hat{m}$, where

$$\hat{m} = \sum_{j=1}^{k-1} x_j.$$

Proof. Let

$$G(q, \lambda) = g(q) + \lambda[1 - (\sum_{j=1}^{k-2} q_j + (3/2)q_{k-1})].$$

We have $\partial G/\partial \lambda = 1 - (\sum_{j=1}^{k-2} q_j + (3/2)q_{k-1})$ and

$$\frac{\partial g}{\partial q_j} = x_j/q_j - \lambda, j = 1, \cdots, k-2; \frac{\partial g}{\partial q_{k-1}} = x_{k-1}/q_{k-1} - (3/2)\lambda.$$

Setting all the partial derivatives to be zero yields

$$q_j \lambda = x_j, j = 1, \cdots, k-2; (3/2)q_{k-1}\lambda = x_{k-1}.$$

Adding all the equations above gives the solution of $\lambda = \hat{m}$ and so the solutions for q_j. The lemma is proved.

This page is intentionally left blank

Global Journal of Science Frontier Research: F
Mathematics and Decision Sciences
Volume 21 Issue 5 Version 1.0 Year 2021
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals
Online ISSN: 2249-4626 & Print ISSN: 0975-5896

A Modification to the Shannon Formula

By J. Ladvánszky

Abstract- Most recently, an experiment was published that cannot be described by the Shannon formula. We will give the details to the reader to check the deviation between the Shannon formula and the experiment. Then we derive a modification of the formula. Consequences of the modification is extended generality, and clear bounds for the average rate of reliable information transfer on a noisy channel of finite bandwidth.

Keywords: *digital communications, noise, shannon formula.*

GJSFR-F Classification: *MSC 2010: 94A24*

Notes

A Modification to the Shannon Formula

J. Ladvánszky

Abstract- Most recently, an experiment was published that cannot be described by the Shannon formula. We will give the details to the reader to check the deviation between the Shannon formula and the experiment. Then we derive a modification of the formula. Consequences of the modification is extended generality, and clear bounds for the average rate of reliable information transfer on a noisy channel of finite bandwidth.

Keywords: digital communications, noise, shannon formula.

I. Introduction

It is a crucial problem in theory of communications if the average speed of information transfer over a noisy channel of finite bandwidth is bounded. That was a seminal result by Shannon (Shannon, 1948) 72 years ago. It is high time, however, to take this result under a magnifying glass, while all kinds of appreciations are given. The long-elapsed time since its discovery proves that the result is robust, it is not easy to find anything to refine.

This almost impossible task is the goal of our paper. Our motivation is to explain the deviation between the Shannon formula and our Matlab/Simulink experiment with a noise insensitive circuit (Ladvánszky, 2020).

II. Problem Statement

In this Section an experiment is described that cannot be characterized by the Shannon formula.

Let us start with the Shannon formula (Shannon, 1948).

$$\mathrm{C} = \mathrm{W}^{*}\log_2\left(1+\frac{\mathrm{S}}{\mathrm{N}}\right) \tag{1}$$

where C is said to be channel capacity, W is the so-called channel bandwidth, S and N are signal and noise average power.

In our Matlab/Simulink experiment (Ladvánszky, 2020), a noisy 4QAM signal is recovered by a modified Costas loop. $W = 4$ MHz, $\frac{S}{N} = -13$ dB, and the 2 MS/s signal is recovered at SER < 0.001 (symbol error rate).

Substituting these data into (1), the result is $C \approx 282$ kS/s, and that roughly deviates from the speed of the recovered signal. As a conclusion, the Shannon formula is not valid for this case.

III. A Possible Solution: A New Formula

In this Section, we gather some facts in connection with the Shannon formula. After some criticism, a new formula is derived that can characterize the experimental fact mentioned in the previous Section.

We start with the classical model of communications, (Shannon, 1948), Figure 1:

Author: *Formerly with Ericsson Hungary, 2013 Pomáz, Vróci u. 3. Hungary. e-mail: ladvanszky55@t-online.hu*

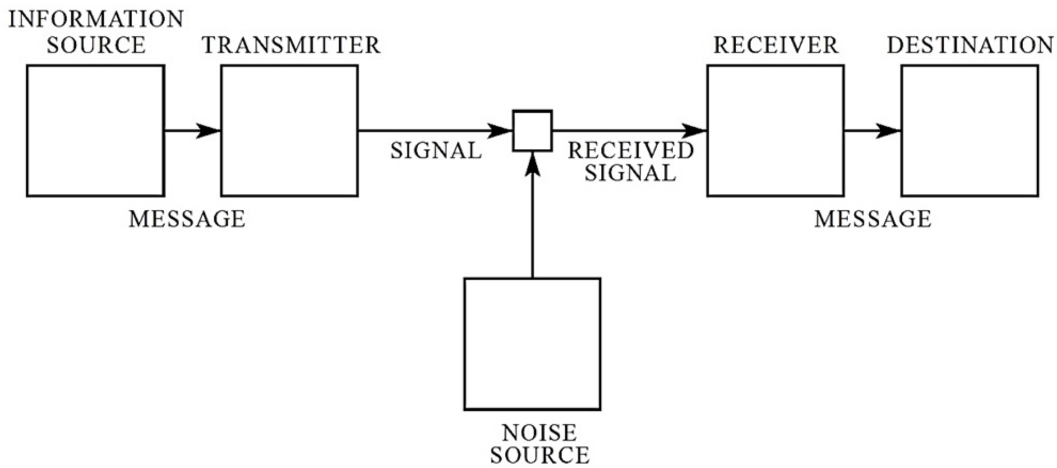

Figure 1: Schematic diagram of a general communication system (Shannon, 1948)

At the center of Figure 1, the small square is the channel. In this paper, we define the channel as a broad-band system block that represents propagation of electromagnetic waves from the transmitter to the receiver. Examples for channel in this sense is a piece of coaxial cable or optical fiber.

The channel in Figure 1 does not contain any filters. Filters are parts of the transmitter and the receiver.

We have three bandwidths in consideration. W is the channel bandwidth, we mean the bandwidth of the channel defined above. In case of coaxial cables, its value is some hundreds of MHz or in the GHz range.

B_S is the signal bandwidth. It is determined by the transmitted signal in the baseband. Its meaning is clarified in Figure 2, its value is 4 MHz in our experiment.

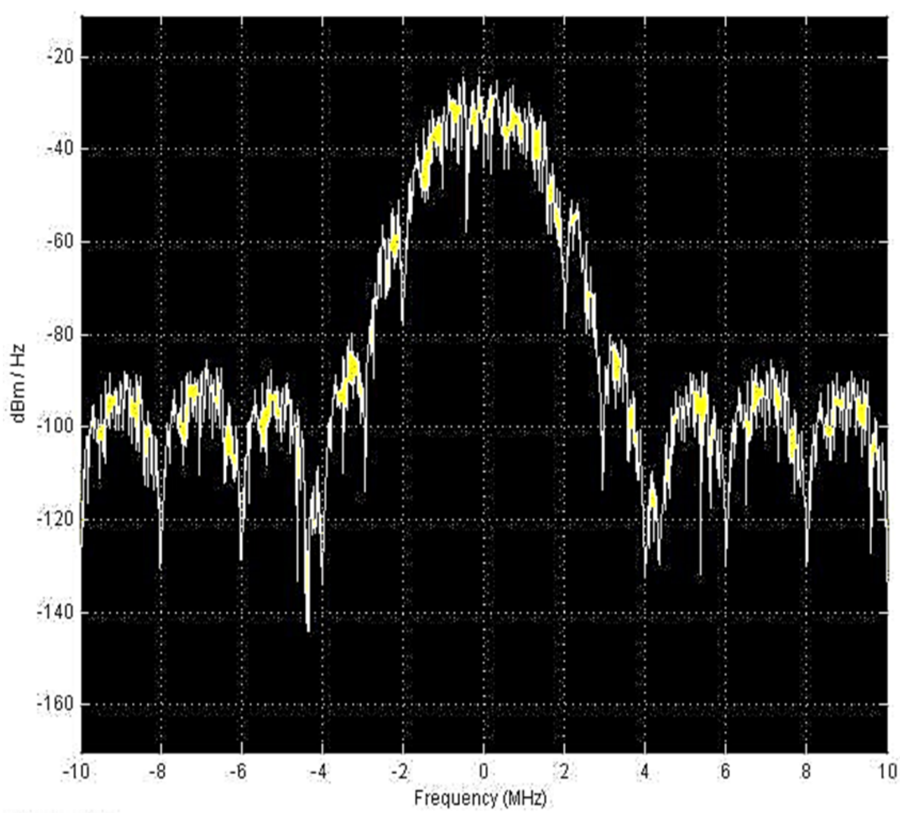

Figure 2: Definition of signal bandwidth on the spectrum of the transmitted baseband signal before transmitter mixer

Notes

$\boldsymbol{B_S} = \mathbf{4}$ MHz, $\boldsymbol{B_N}$ is the noise bandwidth, Figure 3. In our experiment, it is 200 MHz.

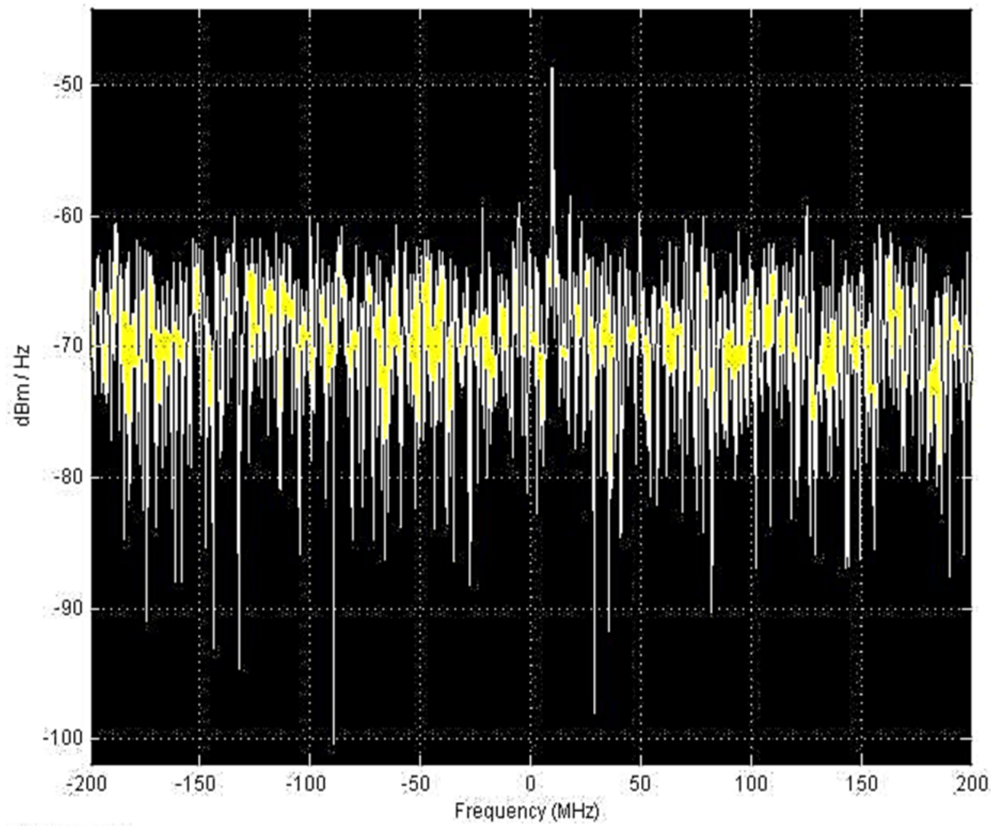

Figure 3: Definition of noise bandwidth on the spectrum of the receiver input signal. $\boldsymbol{B_N} = 200$ MHz

Equation (1) is valid for the case when signal and noise bandwidths are identical (Shannon, 1948).

Our first objection is that C is not a characteristic of the channel alone. For example, S is signal power, a characteristic of the transmitter and not the channel, and W is also questionable. We call here C as upper bound of average speed of information transfer, it is a characteristic of the whole communication system consisting of a transmitter, a channel, and a receiver.

Original explanation for C is (Shannon, 1948), Theorem 12; (Shannon, 1949), (1).

$$\mathrm{C} = \lim_{\mathrm{T}\to\infty} \frac{\log_2 \mathrm{M}}{\mathrm{T}} \tag{2}$$

where $\log_2 M$ is the maximum number of correctly transferred bits. Averaging time is exceptionally long. Name of upper bound for average bitrate is correct.

In (Shannon, 1948), signal bandwidth B_S and noise bandwidth B_N are identical and it is called as channel bandwidth W:

$$\mathrm{B_S} = \mathrm{B_N} = \mathrm{W} \tag{3}$$

where spectra are defined as follows:

$$S(f) = \begin{vmatrix} PDS\ if\ 0 \le f \le B_S \\ 0\ if\ f > B_S \end{vmatrix} \tag{4}$$

$$N(f) = \begin{cases} PDN\ if\ 0 \le f \le B_N \\ 0\ if\ f > B_N \end{cases} \quad (5)$$

where PDS and PDN are power densities, assumed frequency independent.

In the original papers (Shannon, 1948, 1949), the following assumption is hold:

$$B_S = B_N \quad (6)$$

However, in the model of Figure 1, in general,

$$B_S \neq B_N \quad (7)$$

As $B_S > B_N$ is not reasonable, $B_S \le B_N$. Signal to noise ratio is:

$$SNR = \frac{S}{N} \quad (8)$$

Combining (4, 5, 6, 8)

$$SNR = \frac{PDS}{PDN} \quad (9)$$

In general,

$$SNR = \frac{PDS * B_S}{PDN * B_N} \quad (10)$$

If (7) holds, then

$$C \neq B_S * log_2\left(1 + \frac{S}{N}\right) \quad (11)$$

We introduce a new quantity C' That is the formal extension of (1) for the case of non-identical bandwidths:

$$C' = B_S * log_2\left(1 + \frac{S}{N}\right) \quad (1a)$$

In order to extend (1) for the case of non-identical bandwidths (7), let us start from the Shannon formula for colored noise (Shannon, 1949), Equation (32):

$$C_1 = \int_0^{B_S} log_2\left(1 + \frac{S(f)}{N(f)}\right) df \quad (12)$$

By combining (4, 5, 12), we obtain

$$C_1 = B_S * log_2\left(1 + \frac{PDS}{PDN}\right) \quad (13)$$

It is exploited that *PDS* and *PDN* do not depend on f. (13) is valid for the case of non-identical bandwidths (7), while (1) is not valid.

If $B_N > B_S$, then $C_1 > C'$, as we realized it in a Matlab/Simulink experiment with our version of the modified Costas loop (Ladvánszky, 2020).

So, $C_1 > C'$. C' is the formal extension of the Shannon formula, C_1 is our formula. Can we say now if the Shannon formula is exceeded? Author's opinion is yes if also a physical circuit realization is produced. The circuit should operate with low symbol error rate at a SNR negative in dB. This experiment is coming soon.

Notes

IV. Consequences: Bounds for the Average Speed of Communications

Lower bound for operation of the circuit is:

$$\frac{PDS}{PDN} \geq 1 \tag{14}$$

From (12) and (14):

Notes

$$C_1 \geq B_S \tag{15}$$

Average speed of information transfer is bounded from below by the signal bandwidth. That is natural and obvious. Considering (12) for fixed B_S and PDS, we see that C_1 is bounded from above by the smallest possible PDN only. That is the cosmic background radiation (Bucher, 2015) or the thermal noise (Kerr & Randa, 2010), which is greater. At the moment, we can see another limiting factor.

Carrier frequency f_c. In order to be able to reconstruct the carrier,

$$C_1 < f_c \tag{16}$$

This is because the lower edge of the lower sideband around f_c, should be at a positive frequency.

In our 4QAM experiment, actual transfer speed is 2 MS/s and the carrier frequency is 10 MHz.

Now we substitute (16) back into (13):

$$max\left(\frac{PDS}{PDN}\right) = 2^{\frac{f_c}{B_S}} - 1 \tag{17}$$

Meaning of (17) is that a $\frac{PDS}{PDN}$ value greater than that in (17), cannot be exploited for further increasing the bitrate.

For control, we read the power density values from Figure 3. $PDS = -52$ dBm/Hz, $PDN = -72$ dBm/Hz, and we also know that $B_S = 4$ MHz, $f_c = 10$ MHz. Then in (17), $\frac{PDS}{PDN} = 100$ and $2^{\frac{f_c}{B_S}} - 1 \approx 4.66$ so in this case f_c limits the maximum bitrate that can be achieved with increasing PDS.

We can calculate how much it is. Substituting these values in (13), $C_1 \approx 26.6$ Mb/s, and due to the limitation by f_c, $C_1 = 10$ Mb/s could be achieved in ideal case.

V. Conclusions

We discussed some concepts in connection with the Shannon formula. We derived a modification of the original formula and concluded that speed of communications can be higher than that we believed.

If the reader checks the keyword "digital communications pdf" in an internet browser, then it gives more than 20 choices (Proakis & Salehi, 2015). Most of them consider W as channel bandwidth and make signal and noise bandwidths identical. So, our present discussion is timely.

Our derivation is valid if the channel signal is real. The relations are easy to extend for the case of complex signals, however.

It can be clearly recognized that the whole thing depends on the definition of bandwidths. Author's experience is that the typical misunderstanding is as follows. The channel bandwidth is limited to the signal bandwidth and thus a part of the noise spectrum is neglected in calculation of *SNR*. This error occurs if one considers the receiver filter as part of the channel.

We call the attention of the reader to a fine moment in derivation of (13). The same equation can also be achieved when we substitute equal bandwidths (6) in (1a). But what we do is different. We consider different bandwidths (7) and the result in (13) is the consequence of the fact that the upper limit of integration is lower than B_N in (12).

We are aware, by comparison of Figure 2 and Equation (4), that Equation (4) is an approximation. However, considering the exact value of *S(f)* is straightforward.

The author is also available on ResearchGate, LinkedIn and Facebook.

Acknowledgements

The author is grateful to Dr. I. Schmideg who called his attention to check the Shannon formula. Dr. E. Simonyi is also acknowledged for the constant encouragement as the advisor of the DSc work of the author. The Hungarian team representing the medicine NORMALIFE is acknowledged for their significant improvement to the health of the author.

Notes

References Références Referencias

1. Bucher, M. (2015) Physics of the cosmic microwave background anisotropy. *International Journal of Modern Physics D,* arXiv: 1501.04288v1 [astro-ph.CO] (Januay 18th, 2015).
2. Kerr, A. R., & Randa, J. (2010). Thermal Noise and Noise Measurements- A 2010 Update. *IEEE Microwave Magazine*, *11*(6). https://doi.org/10.1109/MMM.2010. 937732
3. Ladvánszky, J. (2020). Noise reduction for digital communications - The Masterpiece, a modified Costas loop. *Circuits and Systems, 11*(6), 57-64. https:// doi.org/10.4236/cs.2020.116006
4. Proakis, J. G., & Salehi, M. (2015). *Digital Communications* (5th ed.). McGraw-Hill, 2008.
5. Shannon, C. E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal, 27*(3), 379-656. https://doi.org/10.1002/j.1538-7305.1948.tb 00917.x
6. Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, *137*, 10-21. https://doi.org/10.1109/JRPROC.1949.232969

Global Journals Guidelines Handbook 2021

www.GlobalJournals.org

Memberships

FELLOWS/ASSOCIATES OF SCIENCE FRONTIER RESEARCH COUNCIL

FSFRC/ASFRC MEMBERSHIPS

Introduction

FSFRC/ASFRC is the most prestigious membership of Global Journals accredited by Open Association of Research Society, U.S.A (OARS). The credentials of Fellow and Associate designations signify that the researcher has gained the knowledge of the fundamental and high-level concepts, and is a subject matter expert, proficient in an expertise course covering the professional code of conduct, and follows recognized standards of practice. The credentials are designated only to the researchers, scientists, and professionals that have been selected by a rigorous process by our Editorial Board and Management Board.

Associates of FSFRC/ASFRC are scientists and researchers from around the world are working on projects/researches that have huge potentials. Members support Global Journals' mission to advance technology for humanity and the profession.

FSFRC

FELLOW OF SCIENCE FRONTIER RESEARCH COUNCIL

FELLOW OF SCIENCE FRONTIER RESEARCH COUNCIL is the most prestigious membership of Global Journals. It is an award and membership granted to individuals that the Open Association of Research Society judges to have made a 'substantial contribution to the improvement of computer science, technology, and electronics engineering.

The primary objective is to recognize the leaders in research and scientific fields of the current era with a global perspective and to create a channel between them and other researchers for better exposure and knowledge sharing. Members are most eminent scientists, engineers, and technologists from all across the world. Fellows are elected for life through a peer review process on the basis of excellence in the respective domain. There is no limit on the number of new nominations made in any year. Each year, the Open Association of Research Society elect up to 12 new Fellow Members.

Benefit

To the institution

Get letter of appreciation

Global Journals sends a letter of appreciation of author to the Dean or CEO of the University or Company of which author is a part, signed by editor in chief or chief author.

Exclusive Network

Get access to a closed network

A FSFRC member gets access to a closed network of Tier 1 researchers and scientists with direct communication channel through our website. Fellows can reach out to other members or researchers directly. They should also be open to reaching out by other.

Career | Credibility | Exclusive | Reputation

Certificate

Receive a print ed copy of a certificate

Fellows receive a printed copy of a certificate signed by our Chief Author that may be used for academic purposes and a personal recommendation letter to the dean of member's university.

Career | Credibility | Exclusive | Reputation

Designation

Get honored title of membership

Fellows can use the honored title of membership. The “FSFRC” is an honored title which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., FSFRC or William Walldroff, M.S., FSFRC.

Career | Credibility | Exclusive | Reputation

Recognition on the Platform

Better visibility and citation

All the Fellow members of FSFRC get a badge of "Leading Member of Global Journals" on the Research Community that distinguishes them from others. Additionally, the profile is also partially maintained by our team for better visibility and citation. All fellows get a dedicated page on the website with their biography.

Career | Credibility | Reputation

Future Work

Get discounts on the future publications

Fellows receive discounts on future publications with Global Journals up to 60%. Through our recommendation programs, members also receive discounts on publications made with OARS affiliated organizations.

Career Financial

GJ Internal Account

Unlimited forward of Emails

Fellows get secure and fast GJ work emails with unlimited forward of emails that they may use them as their primary email. For example,
john [AT] globaljournals [DOT] org.

Career Credibility Reputation

Premium Tools

Access to all the premium tools

To take future researches to the zenith, fellows and associates receive access to all the premium tools that Global Journals have to offer along with the partnership with some of the best marketing leading tools out there.

Financial

Conferences & Events

Organize seminar/conference

Fellows are authorized to organize symposium/seminar/conference on behalf of Global Journal Incorporation (USA). They can also participate in the same organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent. Additionally, they get free research conferences (and others) alerts.

Career Credibility Financial

Early Invitations

Early invitations to all the symposiums, seminars, conferences

All fellows receive the early invitations to all the symposiums, seminars, conferences and webinars hosted by Global Journals in their subject.

Exclusive

Publishing Articles & Books

Earn 60% of sales proceeds

Fellows can publish articles (limited) without any fees. Also, they can earn up to 60% of sales proceeds from the sale of reference/review books/literature/ publishing of research paper. The FSFRC member can decide its price and we can help in making the right decision.

Exclusive Financial

Reviewers

Get a remuneration of 15% of author fees

Fellow members are eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get a remuneration of 15% of author fees, taken from the author of a respective paper.

Financial

Access to Editorial Board

Become a member of the Editorial Board

Fellows may join as a member of the Editorial Board of Global Journals Incorporation (USA) after successful completion of three years as Fellow and as Peer Reviewer. Additionally, Fellows get a chance to nominate other members for Editorial Board.

Career Credibility Exclusive Reputation

And Much More

Get access to scientific museums and observatories across the globe

All members get access to 5 selected scientific museums and observatories across the globe. All researches published with Global Journals will be kept under deep archival facilities across regions for future protections and disaster recovery. They get 10 GB free secure cloud access for storing research files.

ASSOCIATE OF SCIENCE FRONTIER RESEARCH COUNCIL

ASSOCIATE OF SCIENCE FRONTIER RESEARCH COUNCIL is the membership of Global Journals awarded to individuals that the Open Association of Research Society judges to have made a 'substantial contribution to the improvement of computer science, technology, and electronics engineering.

> The primary objective is to recognize the leaders in research and scientific fields of the current era with a global perspective and to create a channel between them and other researchers for better exposure and knowledge sharing. Members are most eminent scientists, engineers, and technologists from all across the world. Associate membership can later be promoted to Fellow Membership. Associates are elected for life through a peer review process on the basis of excellence in the respective domain. There is no limit on the number of new nominations made in any year. Each year, the Open Association of Research Society elect up to 12 new Associate Members.

Benefit

To the institution

Get letter of appreciation

Global Journals sends a letter of appreciation of author to the Dean or CEO of the University or Company of which author is a part, signed by editor in chief or chief author.

Exclusive Network

Get access to a closed network

A ASFRC member gets access to a closed network of Tier 1 researchers and scientists with direct communication channel through our website. Associates can reach out to other members or researchers directly. They should also be open to reaching out by other.

Career | Credibility | Exclusive | Reputation

Certificate

Receive a print ed copy of a certificate

Associates receive a printed copy of a certificate signed by our Chief Author that may be used for academic purposes and a personal recommendation letter to the dean of member's university.

Career | Credibility | Exclusive | Reputation

Designation

Get honored title of membership

Associates can use the honored title of membership. The “ASFRC” is an honored title which is accorded to a person’s name viz. Dr. John E. Hall, Ph.D., ASFRC or William Walldroff, M.S., ASFRC.

Career | Credibility | Exclusive | Reputation

Recognition on the Platform

Better visibility and citation

All the Associate members of ASFRC get a badge of "Leading Member of Global Journals" on the Research Community that distinguishes them from others. Additionally, the profile is also partially maintained by our team for better visibility and citation. All associates get a dedicated page on the website with their biography.

Career | Credibility | Reputation

Future Work

Get discounts on the future publications

Associates receive discounts on the future publications with Global Journals up to 60%. Through our recommendation programs, members also receive discounts on publications made with OARS affiliated organizations.

Career Financial

GJ Internal Account

Unlimited forward of Emails

Associates get secure and fast GJ work emails with unlimited forward of emails that they may use them as their primary email. For example, john [AT] globaljournals [DOT] org.

Career Credibility Reputation

Premium Tools

Access to all the premium tools

To take future researches to the zenith, fellows receive access to almost all the premium tools that Global Journals have to offer along with the partnership with some of the best marketing leading tools out there.

Financial

Conferences & Events

Organize seminar/conference

Associates are authorized to organize symposium/seminar/conference on behalf of Global Journal Incorporation (USA). They can also participate in the same organized by another institution as representative of Global Journal. In both the cases, it is mandatory for him to discuss with us and obtain our consent. Additionally, they get free research conferences (and others) alerts.

Career Credibility Financial

Early Invitations

Early invitations to all the symposiums, seminars, conferences

All associates receive the early invitations to all the symposiums, seminars, conferences and webinars hosted by Global Journals in their subject.

Exclusive

Publishing Articles & Books

Earn 30-40% of sales proceeds

Associates can publish articles (limited) without any fees. Also, they can earn up to 30-40% of sales proceeds from the sale of reference/review books/literature/publishing of research paper.

Exclusive Financial

Reviewers

Get a remuneration of 15% of author fees

Associate members are eligible to join as a paid peer reviewer at Global Journals Incorporation (USA) and can get a remuneration of 15% of author fees, taken from the author of a respective paper.

Financial

And Much More

Get access to scientific museums and observatories across the globe

All members get access to 2 selected scientific museums and observatories across the globe. All researches published with Global Journals will be kept under deep archival facilities across regions for future protections and disaster recovery. They get 5 GB free secure cloud access for storing research files.

Associate	Fellow	Research Group	Basic
$4800 **lifetime designation**	$6800 **lifetime designation**	$12500.00 **organizational**	APC **per article**
Certificate, LoR and Momento 2 discounted publishing/year **Gradation** of Research 10 research contacts/day 1 **GB** Cloud Storage **GJ** Community Access	**Certificate,** LoR and Momento **Unlimited** discounted publishing/year **Gradation** of Research **Unlimited** research contacts/day 5 **GB** Cloud Storage **Online Presense** Assistance **GJ** Community Access	**Certificates,** LoRs and Momentos **Unlimited** free publishing/year **Gradation** of Research **Unlimited** research contacts/day **Unlimited** Cloud Storage **Online Presense** Assistance **GJ** Community Access	**GJ** Community Access

Preferred Author Guidelines

We accept the manuscript submissions in any standard (generic) format.

We typeset manuscripts using advanced typesetting tools like Adobe In Design, CorelDraw, TeXnicCenter, and TeXStudio. We usually recommend authors submit their research using any standard format they are comfortable with, and let Global Journals do the rest.

Alternatively, you can download our basic template from https://globaljournals.org/Template.zip

Authors should submit their complete paper/article, including text illustrations, graphics, conclusions, artwork, and tables. Authors who are not able to submit manuscript using the form above can email the manuscript department at submit@globaljournals.org or get in touch with chiefeditor@globaljournals.org if they wish to send the abstract before submission.

Before and during Submission

Authors must ensure the information provided during the submission of a paper is authentic. Please go through the following checklist before submitting:

1. Authors must go through the complete author guideline and understand and *agree to Global Journals' ethics and code of conduct,* along with author responsibilities.
2. Authors must accept the privacy policy, terms, and conditions of Global Journals.
3. Ensure corresponding author's email address and postal address are accurate and reachable.
4. Manuscript to be submitted must include keywords, an abstract, a paper title, co-author(s') names and details (email address, name, phone number, and institution), figures and illustrations in vector format including appropriate captions, tables, including titles and footnotes, a conclusion, results, acknowledgments and references.
5. Authors should submit paper in a ZIP archive if any supplementary files are required along with the paper.
6. Proper permissions must be acquired for the use of any copyrighted material.
7. Manuscript submitted *must not have been submitted or published elsewhere* and all authors must be aware of the submission.

Declaration of Conflicts of Interest

It is required for authors to declare all financial, institutional, and personal relationships with other individuals and organizations that could influence (bias) their research.

Policy on Plagiarism

Plagiarism is not acceptable in Global Journals submissions at all.

Plagiarized content will not be considered for publication. We reserve the right to inform authors' institutions about plagiarism detected either before or after publication. If plagiarism is identified, we will follow COPE guidelines:

Authors are solely responsible for all the plagiarism that is found. The author must not fabricate, falsify or plagiarize existing research data. The following, if copied, will be considered plagiarism:

- Words (language)
- Ideas
- Findings
- Writings
- Diagrams
- Graphs
- Illustrations
- Lectures

- Printed material
- Graphic representations
- Computer programs
- Electronic material
- Any other original work

Authorship Policies

Global Journals follows the definition of authorship set up by the Open Association of Research Society, USA. According to its guidelines, authorship criteria must be based on:

1. Substantial contributions to the conception and acquisition of data, analysis, and interpretation of findings.
2. Drafting the paper and revising it critically regarding important academic content.
3. Final approval of the version of the paper to be published.

Changes in Authorship

The corresponding author should mention the name and complete details of all co-authors during submission and in manuscript. We support addition, rearrangement, manipulation, and deletions in authors list till the early view publication of the journal. We expect that corresponding author will notify all co-authors of submission. We follow COPE guidelines for changes in authorship.

Copyright

During submission of the manuscript, the author is confirming an exclusive license agreement with Global Journals which gives Global Journals the authority to reproduce, reuse, and republish authors' research. We also believe in flexible copyright terms where copyright may remain with authors/employers/institutions as well. Contact your editor after acceptance to choose your copyright policy. You may follow this form for copyright transfers.

Appealing Decisions

Unless specified in the notification, the Editorial Board's decision on publication of the paper is final and cannot be appealed before making the major change in the manuscript.

Acknowledgments

Contributors to the research other than authors credited should be mentioned in Acknowledgments. The source of funding for the research can be included. Suppliers of resources may be mentioned along with their addresses.

Declaration of funding sources

Global Journals is in partnership with various universities, laboratories, and other institutions worldwide in the research domain. Authors are requested to disclose their source of funding during every stage of their research, such as making analysis, performing laboratory operations, computing data, and using institutional resources, from writing an article to its submission. This will also help authors to get reimbursements by requesting an open access publication letter from Global Journals and submitting to the respective funding source.

Preparing your Manuscript

Authors can submit papers and articles in an acceptable file format: MS Word (doc, docx), LaTeX (.tex, .zip or .rar including all of your files), Adobe PDF (.pdf), rich text format (.rtf), simple text document (.txt), Open Document Text (.odt), and Apple Pages (.pages). Our professional layout editors will format the entire paper according to our official guidelines. This is one of the highlights of publishing with Global Journals—authors should not be concerned about the formatting of their paper. Global Journals accepts articles and manuscripts in every major language, be it Spanish, Chinese, Japanese, Portuguese, Russian, French, German, Dutch, Italian, Greek, or any other national language, but the title, subtitle, and abstract should be in English. This will facilitate indexing and the pre-peer review process.

The following is the official style and template developed for publication of a research paper. Authors are not required to follow this style during the submission of the paper. It is just for reference purposes.

Manuscript Style Instruction (Optional)

- Microsoft Word Document Setting Instructions.
- Font type of all text should be Swis721 Lt BT.
- Page size: 8.27" x 11'", left margin: 0.65, right margin: 0.65, bottom margin: 0.75.
- Paper title should be in one column of font size 24.
- Author name in font size of 11 in one column.
- Abstract: font size 9 with the word "Abstract" in bold italics.
- Main text: font size 10 with two justified columns.
- Two columns with equal column width of 3.38 and spacing of 0.2.
- First character must be three lines drop-capped.
- The paragraph before spacing of 1 pt and after of 0 pt.
- Line spacing of 1 pt.
- Large images must be in one column.
- The names of first main headings (Heading 1) must be in Roman font, capital letters, and font size of 10.
- The names of second main headings (Heading 2) must not include numbers and must be in italics with a font size of 10.

Structure and Format of Manuscript

The recommended size of an original research paper is under 15,000 words and review papers under 7,000 words. Research articles should be less than 10,000 words. Research papers are usually longer than review papers. Review papers are reports of significant research (typically less than 7,000 words, including tables, figures, and references)

A research paper must include:

a) A title which should be relevant to the theme of the paper.
b) A summary, known as an abstract (less than 150 words), containing the major results and conclusions.
c) Up to 10 keywords that precisely identify the paper's subject, purpose, and focus.
d) An introduction, giving fundamental background objectives.
e) Resources and techniques with sufficient complete experimental details (wherever possible by reference) to permit repetition, sources of information must be given, and numerical methods must be specified by reference.
f) Results which should be presented concisely by well-designed tables and figures.
g) Suitable statistical data should also be given.
h) All data must have been gathered with attention to numerical detail in the planning stage.

Design has been recognized to be essential to experiments for a considerable time, and the editor has decided that any paper that appears not to have adequate numerical treatments of the data will be returned unrefereed.

i) Discussion should cover implications and consequences and not just recapitulate the results; conclusions should also be summarized.
j) There should be brief acknowledgments.
k) There ought to be references in the conventional format. Global Journals recommends APA format.

Authors should carefully consider the preparation of papers to ensure that they communicate effectively. Papers are much more likely to be accepted if they are carefully designed and laid out, contain few or no errors, are summarizing, and follow instructions. They will also be published with much fewer delays than those that require much technical and editorial correction.

The Editorial Board reserves the right to make literary corrections and suggestions to improve brevity.

Format Structure

It is necessary that authors take care in submitting a manuscript that is written in simple language and adheres to published guidelines.

All manuscripts submitted to Global Journals should include:

Title

The title page must carry an informative title that reflects the content, a running title (less than 45 characters together with spaces), names of the authors and co-authors, and the place(s) where the work was carried out.

Author details

The full postal address of any related author(s) must be specified.

Abstract

The abstract is the foundation of the research paper. It should be clear and concise and must contain the objective of the paper and inferences drawn. It is advised to not include big mathematical equations or complicated jargon.

Many researchers searching for information online will use search engines such as Google, Yahoo or others. By optimizing your paper for search engines, you will amplify the chance of someone finding it. In turn, this will make it more likely to be viewed and cited in further works. Global Journals has compiled these guidelines to facilitate you to maximize the web-friendliness of the most public part of your paper.

Keywords

A major lynchpin of research work for the writing of research papers is the keyword search, which one will employ to find both library and internet resources. Up to eleven keywords or very brief phrases have to be given to help data retrieval, mining, and indexing.

One must be persistent and creative in using keywords. An effective keyword search requires a strategy: planning of a list of possible keywords and phrases to try.

Choice of the main keywords is the first tool of writing a research paper. Research paper writing is an art. Keyword search should be as strategic as possible.

One should start brainstorming lists of potential keywords before even beginning searching. Think about the most important concepts related to research work. Ask, "What words would a source have to include to be truly valuable in a research paper?" Then consider synonyms for the important words.

It may take the discovery of only one important paper to steer in the right keyword direction because, in most databases, the keywords under which a research paper is abstracted are listed with the paper.

Numerical Methods

Numerical methods used should be transparent and, where appropriate, supported by references.

Abbreviations

Authors must list all the abbreviations used in the paper at the end of the paper or in a separate table before using them.

Formulas and equations

Authors are advised to submit any mathematical equation using either MathJax, KaTeX, or LaTeX, or in a very high-quality image.

Tables, Figures, and Figure Legends

Tables: Tables should be cautiously designed, uncrowned, and include only essential data. Each must have an Arabic number, e.g., Table 4, a self-explanatory caption, and be on a separate sheet. Authors must submit tables in an editable format and not as images. References to these tables (if any) must be mentioned accurately.

Figures

Figures are supposed to be submitted as separate files. Always include a citation in the text for each figure using Arabic numbers, e.g., Fig. 4. Artwork must be submitted online in vector electronic form or by emailing it.

Preparation of Eletronic Figures for Publication

Although low-quality images are sufficient for review purposes, print publication requires high-quality images to prevent the final product being blurred or fuzzy. Submit (possibly by e-mail) EPS (line art) or TIFF (halftone/ photographs) files only. MS PowerPoint and Word Graphics are unsuitable for printed pictures. Avoid using pixel-oriented software. Scans (TIFF only) should have a resolution of at least 350 dpi (halftone) or 700 to 1100 dpi (line drawings). Please give the data for figures in black and white or submit a Color Work Agreement form. EPS files must be saved with fonts embedded (and with a TIFF preview, if possible).

For scanned images, the scanning resolution at final image size ought to be as follows to ensure good reproduction: line art: >650 dpi; halftones (including gel photographs): >350 dpi; figures containing both halftone and line images: >650 dpi.

Color charges: Authors are advised to pay the full cost for the reproduction of their color artwork. Hence, please note that if there is color artwork in your manuscript when it is accepted for publication, we would require you to complete and return a Color Work Agreement form before your paper can be published. Also, you can email your editor to remove the color fee after acceptance of the paper.

Tips for Writing a Good Quality Science Frontier Research Paper

Techniques for writing a good quality Science Frontier Research paper:

1. Choosing the topic: In most cases, the topic is selected by the interests of the author, but it can also be suggested by the guides. You can have several topics, and then judge which you are most comfortable with. This may be done by asking several questions of yourself, like "Will I be able to carry out a search in this area? Will I find all necessary resources to accomplish the search? Will I be able to find all information in this field area?" If the answer to this type of question is "yes," then you ought to choose that topic. In most cases, you may have to conduct surveys and visit several places. Also, you might have to do a lot of work to find all the rises and falls of the various data on that subject. Sometimes, detailed information plays a vital role, instead of short information. Evaluators are human: The first thing to remember is that evaluators are also human beings. They are not only meant for rejecting a paper. They are here to evaluate your paper. So present your best aspect.

2. Think like evaluators: If you are in confusion or getting demotivated because your paper may not be accepted by the evaluators, then think, and try to evaluate your paper like an evaluator. Try to understand what an evaluator wants in your research paper, and you will automatically have your answer. Make blueprints of paper: The outline is the plan or framework that will help you to arrange your thoughts. It will make your paper logical. But remember that all points of your outline must be related to the topic you have chosen.

3. Ask your guides: If you are having any difficulty with your research, then do not hesitate to share your difficulty with your guide (if you have one). They will surely help you out and resolve your doubts. If you can't clarify what exactly you require for your work, then ask your supervisor to help you with an alternative. He or she might also provide you with a list of essential readings.

4. Use of computer is recommended: As you are doing research in the field of science frontier then this point is quite obvious. Use right software: Always use good quality software packages. If you are not capable of judging good software, then you can lose the quality of your paper unknowingly. There are various programs available to help you which you can get through the internet.

5. Use the internet for help: An excellent start for your paper is using Google. It is a wondrous search engine, where you can have your doubts resolved. You may also read some answers for the frequent question of how to write your research paper or find a model research paper. You can download books from the internet. If you have all the required books, place importance on reading, selecting, and analyzing the specified information. Then sketch out your research paper. Use big pictures: You may use encyclopedias like Wikipedia to get pictures with the best resolution. At Global Journals, you should strictly follow here.

6. Bookmarks are useful: When you read any book or magazine, you generally use bookmarks, right? It is a good habit which helps to not lose your continuity. You should always use bookmarks while searching on the internet also, which will make your search easier.

7. Revise what you wrote: When you write anything, always read it, summarize it, and then finalize it.

8. Make every effort: Make every effort to mention what you are going to write in your paper. That means always have a good start. Try to mention everything in the introduction—what is the need for a particular research paper. Polish your work with good writing skills and always give an evaluator what he wants. Make backups: When you are going to do any important thing like making a research paper, you should always have backup copies of it either on your computer or on paper. This protects you from losing any portion of your important data.

9. Produce good diagrams of your own: Always try to include good charts or diagrams in your paper to improve quality. Using several unnecessary diagrams will degrade the quality of your paper by creating a hodgepodge. So always try to include diagrams which were made by you to improve the readability of your paper. Use of direct quotes: When you do research relevant to literature, history, or current affairs, then use of quotes becomes essential, but if the study is relevant to science, use of quotes is not preferable.

10. Use proper verb tense: Use proper verb tenses in your paper. Use past tense to present those events that have happened. Use present tense to indicate events that are going on. Use future tense to indicate events that will happen in the future. Use of wrong tenses will confuse the evaluator. Avoid sentences that are incomplete.

11. Pick a good study spot: Always try to pick a spot for your research which is quiet. Not every spot is good for studying.

12. Know what you know: Always try to know what you know by making objectives, otherwise you will be confused and unable to achieve your target.

13. Use good grammar: Always use good grammar and words that will have a positive impact on the evaluator; use of good vocabulary does not mean using tough words which the evaluator has to find in a dictionary. Do not fragment sentences. Eliminate one-word sentences. Do not ever use a big word when a smaller one would suffice.

Verbs have to be in agreement with their subjects. In a research paper, do not start sentences with conjunctions or finish them with prepositions. When writing formally, it is advisable to never split an infinitive because someone will (wrongly) complain. Avoid clichés like a disease. Always shun irritating alliteration. Use language which is simple and straightforward. Put together a neat summary.

14. Arrangement of information: Each section of the main body should start with an opening sentence, and there should be a changeover at the end of the section. Give only valid and powerful arguments for your topic. You may also maintain your arguments with records.

15. Never start at the last minute: Always allow enough time for research work. Leaving everything to the last minute will degrade your paper and spoil your work.

16. Multitasking in research is not good: Doing several things at the same time is a bad habit in the case of research activity. Research is an area where everything has a particular time slot. Divide your research work into parts, and do a particular part in a particular time slot.

17. Never copy others' work: Never copy others' work and give it your name because if the evaluator has seen it anywhere, you will be in trouble. Take proper rest and food: No matter how many hours you spend on your research activity, if you are not taking care of your health, then all your efforts will have been in vain. For quality research, take proper rest and food.

18. Go to seminars: Attend seminars if the topic is relevant to your research area. Utilize all your resources.

19. Refresh your mind after intervals: Try to give your mind a rest by listening to soft music or sleeping in intervals. This will also improve your memory. Acquire colleagues: Always try to acquire colleagues. No matter how sharp you are, if you acquire colleagues, they can give you ideas which will be helpful to your research.

20. Think technically: Always think technically. If anything happens, search for its reasons, benefits, and demerits. Think and then print: When you go to print your paper, check that tables are not split, headings are not detached from their descriptions, and page sequence is maintained.

21. Adding unnecessary information: Do not add unnecessary information like "I have used MS Excel to draw graphs." Irrelevant and inappropriate material is superfluous. Foreign terminology and phrases are not apropos. One should never take a broad view. Analogy is like feathers on a snake. Use words properly, regardless of how others use them. Remove quotations. Puns are for kids, not grunt readers. Never oversimplify: When adding material to your research paper, never go for oversimplification; this will definitely irritate the evaluator. Be specific. Never use rhythmic redundancies. Contractions shouldn't be used in a research paper. Comparisons are as terrible as clichés. Give up ampersands, abbreviations, and so on. Remove commas that are not necessary. Parenthetical words should be between brackets or commas. Understatement is always the best way to put forward earth-shaking thoughts. Give a detailed literary review.

22. Report concluded results: Use concluded results. From raw data, filter the results, and then conclude your studies based on measurements and observations taken. An appropriate number of decimal places should be used. Parenthetical remarks are prohibited here. Proofread carefully at the final stage. At the end, give an outline to your arguments. Spot perspectives of further study of the subject. Justify your conclusion at the bottom sufficiently, which will probably include examples.

23. Upon conclusion: Once you have concluded your research, the next most important step is to present your findings. Presentation is extremely important as it is the definite medium though which your research is going to be in print for the rest of the crowd. Care should be taken to categorize your thoughts well and present them in a logical and neat manner. A good quality research paper format is essential because it serves to highlight your research paper and bring to light all necessary aspects of your research.

Informal Guidelines of Research Paper Writing

Key points to remember:

- Submit all work in its final form.
- Write your paper in the form which is presented in the guidelines using the template.
- Please note the criteria peer reviewers will use for grading the final paper.

Final points:

One purpose of organizing a research paper is to let people interpret your efforts selectively. The journal requires the following sections, submitted in the order listed, with each section starting on a new page:

The introduction: This will be compiled from reference matter and reflect the design processes or outline of basis that directed you to make a study. As you carry out the process of study, the method and process section will be constructed like that. The results segment will show related statistics in nearly sequential order and direct reviewers to similar intellectual paths throughout the data that you gathered to carry out your study.

The discussion section:

This will provide understanding of the data and projections as to the implications of the results. The use of good quality references throughout the paper will give the effort trustworthiness by representing an alertness to prior workings.

Writing a research paper is not an easy job, no matter how trouble-free the actual research or concept. Practice, excellent preparation, and controlled record-keeping are the only means to make straightforward progression.

General style:

Specific editorial column necessities for compliance of a manuscript will always take over from directions in these general guidelines.

To make a paper clear: Adhere to recommended page limits.

Mistakes to avoid:

- Insertion of a title at the foot of a page with subsequent text on the next page.
- Separating a table, chart, or figure—confine each to a single page.
- Submitting a manuscript with pages out of sequence.
- In every section of your document, use standard writing style, including articles ("a" and "the").
- Keep paying attention to the topic of the paper.
- Use paragraphs to split each significant point (excluding the abstract).
- Align the primary line of each section.
- Present your points in sound order.
- Use present tense to report well-accepted matters.
- Use past tense to describe specific results.
- Do not use familiar wording; don't address the reviewer directly. Don't use slang or superlatives.
- Avoid use of extra pictures—include only those figures essential to presenting results.

Title page:

Choose a revealing title. It should be short and include the name(s) and address(es) of all authors. It should not have acronyms or abbreviations or exceed two printed lines.

Abstract: This summary should be two hundred words or less. It should clearly and briefly explain the key findings reported in the manuscript and must have precise statistics. It should not have acronyms or abbreviations. It should be logical in itself. Do not cite references at this point.

An abstract is a brief, distinct paragraph summary of finished work or work in development. In a minute or less, a reviewer can be taught the foundation behind the study, common approaches to the problem, relevant results, and significant conclusions or new questions.

Write your summary when your paper is completed because how can you write the summary of anything which is not yet written? Wealth of terminology is very essential in abstract. Use comprehensive sentences, and do not sacrifice readability for brevity; you can maintain it succinctly by phrasing sentences so that they provide more than a lone rationale. The author can at this moment go straight to shortening the outcome. Sum up the study with the subsequent elements in any summary. Try to limit the initial two items to no more than one line each.

Reason for writing the article—theory, overall issue, purpose.

- Fundamental goal.
- To-the-point depiction of the research.
- Consequences, including definite statistics—if the consequences are quantitative in nature, account for this; results of any numerical analysis should be reported. Significant conclusions or questions that emerge from the research.

Approach:

- o Single section and succinct.
- o An outline of the job done is always written in past tense.
- o Concentrate on shortening results—limit background information to a verdict or two.
- o Exact spelling, clarity of sentences and phrases, and appropriate reporting of quantities (proper units, important statistics) are just as significant in an abstract as they are anywhere else.

Introduction:

The introduction should "introduce" the manuscript. The reviewer should be presented with sufficient background information to be capable of comprehending and calculating the purpose of your study without having to refer to other works. The basis for the study should be offered. Give the most important references, but avoid making a comprehensive appraisal of the topic. Describe the problem visibly. If the problem is not acknowledged in a logical, reasonable way, the reviewer will give no attention to your results. Speak in common terms about techniques used to explain the problem, if needed, but do not present any particulars about the protocols here.

The following approach can create a valuable beginning:

- Explain the value (significance) of the study.
- Defend the model—why did you employ this particular system or method? What is its compensation? Remark upon its appropriateness from an abstract point of view as well as pointing out sensible reasons for using it.
- Present a justification. State your particular theory(-ies) or aim(s), and describe the logic that led you to choose them.
- Briefly explain the study's tentative purpose and how it meets the declared objectives.

Approach:

Use past tense except for when referring to recognized facts. After all, the manuscript will be submitted after the entire job is done. Sort out your thoughts; manufacture one key point for every section. If you make the four points listed above, you will need at least four paragraphs. Present surrounding information only when it is necessary to support a situation. The reviewer does not desire to read everything you know about a topic. Shape the theory specifically—do not take a broad view.

As always, give awareness to spelling, simplicity, and correctness of sentences and phrases.

Procedures (methods and materials):

This part is supposed to be the easiest to carve if you have good skills. A soundly written procedures segment allows a capable scientist to replicate your results. Present precise information about your supplies. The suppliers and clarity of reagents can be helpful bits of information. Present methods in sequential order, but linked methodologies can be grouped as a segment. Be concise when relating the protocols. Attempt to give the least amount of information that would permit another capable scientist to replicate your outcome, but be cautious that vital information is integrated. The use of subheadings is suggested and ought to be synchronized with the results section.

When a technique is used that has been well-described in another section, mention the specific item describing the way, but draw the basic principle while stating the situation. The purpose is to show all particular resources and broad procedures so that another person may use some or all of the methods in one more study or referee the scientific value of your work. It is not to be a step-by-step report of the whole thing you did, nor is a methods section a set of orders.

Materials:

Materials may be reported in part of a section or else they may be recognized along with your measures.

Methods:

- Report the method and not the particulars of each process that engaged the same methodology.
- Describe the method entirely.
- To be succinct, present methods under headings dedicated to specific dealings or groups of measures.
- Simplify—detail how procedures were completed, not how they were performed on a particular day.
- If well-known procedures were used, account for the procedure by name, possibly with a reference, and that's all.

Approach:

It is embarrassing to use vigorous voice when documenting methods without using first person, which would focus the reviewer's interest on the researcher rather than the job. As a result, when writing up the methods, most authors use third person passive voice.

Use standard style in this and every other part of the paper—avoid familiar lists, and use full sentences.

What to keep away from:

- Resources and methods are not a set of information.
- Skip all descriptive information and surroundings—save it for the argument.
- Leave out information that is immaterial to a third party.

Results:

The principle of a results segment is to present and demonstrate your conclusion. Create this part as entirely objective details of the outcome, and save all understanding for the discussion.

The page length of this segment is set by the sum and types of data to be reported. Use statistics and tables, if suitable, to present consequences most efficiently.

You must clearly differentiate material which would usually be incorporated in a study editorial from any unprocessed data or additional appendix matter that would not be available. In fact, such matters should not be submitted at all except if requested by the instructor.

Content:

- Sum up your conclusions in text and demonstrate them, if suitable, with figures and tables.
- In the manuscript, explain each of your consequences, and point the reader to remarks that are most appropriate.
- Present a background, such as by describing the question that was addressed by creation of an exacting study.
- Explain results of control experiments and give remarks that are not accessible in a prescribed figure or table, if appropriate.
- Examine your data, then prepare the analyzed (transformed) data in the form of a figure (graph), table, or manuscript.

What to stay away from:

- Do not discuss or infer your outcome, report surrounding information, or try to explain anything.
- Do not include raw data or intermediate calculations in a research manuscript.
- Do not present similar data more than once.
- A manuscript should complement any figures or tables, not duplicate information.
- Never confuse figures with tables—there is a difference.

Approach:

As always, use past tense when you submit your results, and put the whole thing in a reasonable order.

Put figures and tables, appropriately numbered, in order at the end of the report.

If you desire, you may place your figures and tables properly within the text of your results section.

Figures and tables:

If you put figures and tables at the end of some details, make certain that they are visibly distinguished from any attached appendix materials, such as raw facts. Whatever the position, each table must be titled, numbered one after the other, and include a heading. All figures and tables must be divided from the text.

Discussion:

The discussion is expected to be the trickiest segment to write. A lot of papers submitted to the journal are discarded based on problems with the discussion. There is no rule for how long an argument should be.

Position your understanding of the outcome visibly to lead the reviewer through your conclusions, and then finish the paper with a summing up of the implications of the study. The purpose here is to offer an understanding of your results and support all of your conclusions, using facts from your research and generally accepted information, if suitable. The implication of results should be fully described.

Infer your data in the conversation in suitable depth. This means that when you clarify an observable fact, you must explain mechanisms that may account for the observation. If your results vary from your prospect, make clear why that may have happened. If your results agree, then explain the theory that the proof supported. It is never suitable to just state that the data approved the prospect, and let it drop at that. Make a decision as to whether each premise is supported or discarded or if you cannot make a conclusion with assurance. Do not just dismiss a study or part of a study as "uncertain."

Research papers are not acknowledged if the work is imperfect. Draw what conclusions you can based upon the results that you have, and take care of the study as a finished work.

- o You may propose future guidelines, such as how an experiment might be personalized to accomplish a new idea.
- o Give details of all of your remarks as much as possible, focusing on mechanisms.
- o Make a decision as to whether the tentative design sufficiently addressed the theory and whether or not it was correctly restricted. Try to present substitute explanations if they are sensible alternatives.
- o One piece of research will not counter an overall question, so maintain the large picture in mind. Where do you go next? The best studies unlock new avenues of study. What questions remain?
- o Recommendations for detailed papers will offer supplementary suggestions.

Approach:

When you refer to information, differentiate data generated by your own studies from other available information. Present work done by specific persons (including you) in past tense.

Describe generally acknowledged facts and main beliefs in present tense.

The Administration Rules

Administration Rules to Be Strictly Followed before Submitting Your Research Paper to Global Journals Inc.

Please read the following rules and regulations carefully before submitting your research paper to Global Journals Inc. to avoid rejection.

Segment draft and final research paper: You have to strictly follow the template of a research paper, failing which your paper may get rejected. You are expected to write each part of the paper wholly on your own. The peer reviewers need to identify your own perspective of the concepts in your own terms. Please do not extract straight from any other source, and do not rephrase someone else's analysis. Do not allow anyone else to proofread your manuscript.

Written material: You may discuss this with your guides and key sources. Do not copy anyone else's paper, even if this is only imitation, otherwise it will be rejected on the grounds of plagiarism, which is illegal. Various methods to avoid plagiarism are strictly applied by us to every paper, and, if found guilty, you may be blacklisted, which could affect your career adversely. To guard yourself and others from possible illegal use, please do not permit anyone to use or even read your paper and file.

Criterion for Grading a Research Paper (Compilation)
by Global Journals

Please note that following table is only a Grading of "Paper Compilation" and not on "Performed/Stated Research" whose grading solely depends on Individual Assigned Peer Reviewer and Editorial Board Member. These can be available only on request and after decision of Paper. This report will be the property of Global Journals.

Topics	Grades		
	A-B	**C-D**	**E-F**
Abstract	Clear and concise with appropriate content, Correct format. 200 words or below	Unclear summary and no specific data, Incorrect form Above 200 words	No specific data with ambiguous information Above 250 words
Introduction	Containing all background details with clear goal and appropriate details, flow specification, no grammar and spelling mistake, well organized sentence and paragraph, reference cited	Unclear and confusing data, appropriate format, grammar and spelling errors with unorganized matter	Out of place depth and content, hazy format
Methods and Procedures	Clear and to the point with well arranged paragraph, precision and accuracy of facts and figures, well organized subheads	Difficult to comprehend with embarrassed text, too much explanation but completed	Incorrect and unorganized structure with hazy meaning
Result	Well organized, Clear and specific, Correct units with precision, correct data, well structuring of paragraph, no grammar and spelling mistake	Complete and embarrassed text, difficult to comprehend	Irregular format with wrong facts and figures
Discussion	Well organized, meaningful specification, sound conclusion, logical and concise explanation, highly structured paragraph reference cited	Wordy, unclear conclusion, spurious	Conclusion is not cited, unorganized, difficult to comprehend
References	Complete and correct format, well organized	Beside the point, Incomplete	Wrong format and structuring

Index

save our planet

Global Journal of Science Frontier Research

Visit us on the Web at www.GlobalJournals.org | www.JournalofScience.org
or email us at helpdesk@globaljournals.org

ISSN 9755896