International Journal of Research in Medical Sciences
Shanmugam R. Int J Res Med Sci. 2015 Jun;3(6):1336-1341
www.msjonline.org
pISSN 2320-6071 | eISSN 2320-6012

DOI: 10.18203/2320-6012.ijrms20150142

Research Article

Never, once, and repeated illness: a geometric view for insights and interpretations

Ramalingam Shanmugam*

School of Health Administration, Texas State University, San Marcos, TX 78666, USA

Received: 25 March 2015
Accepted: 07 May 2015

***Correspondence:**
Dr. Ramalingam Shanmugam,
E-mail: rs25@txstate.edu, shanmugam@txstate.edu

ABSTRACT

Background: Medical/health researchers depend on data evidence for knowledge discovery. At times, data analysis to capture the data evidence is overwhelming and the process becomes too tedious to give up the attempt. A prudent thing to do is to seek out a simpler visual approach to obtain insights. One visual approach is devised in this article to understand what the data are really revealing to either get an insight first or then confirm what is intuitively configured by the medical concepts. This visual approach is geometric concepts based. In specific, triangle is employed in this new and novel approach.
Methods: A successful treatment of any illness is a consequence of knowledge build-up arising from data mining about the never, once, or repeated episode of a disease incidence in a patient. This article investigates and illustrates a novel and pioneering geometric approach, especially based on the properties of triangle, to extract hidden evidence in the data. New probabilistic expressions are derived utilizing trigonometric relations among the corner points of a triangle. The conceptual contents of this article are versatile enough for different medical/health data analysis.
Results: For illustration here, the medical binomial data in Hopper et al. (Genetic Epidemiology, 1990) on the occurrence of asthma or hay fever among the four groups: (1) monozygotic females (MZF), (2) monozygotic males (MZM), (3) di-zygotic females (DZF), and (4) di-zygotic males (DZM) are considered and triangularly interpreted. The results indicate that the angle in the vertex representing one episode is the largest compared to the other two angles in the vertices representing never or repeated episode of an illness among a random sample of twins from these four groups with respect to getting asthma or hay fever. This geometric finding implies that the event of never and the event of repeated incidence of the illness have farthest Euclidean distance in probability sense. In other words, the never and repeated incidences are not in close proximity as probable.
Conclusions: This geometric view of this article is versatile enough to be useful in other research studies in drug assessment, clinical trial outcomes, business, marketing, finance, economics, engineering and public health whether the data are Poisson or inverse binomial type as well.

Keywords: Discrete probability, Likelihood of two or more episodes, Poisson, Binomial, Geometric models

INTRODUCTION

Geometry is the foundation of every aspect of life. The natural world seems to be regulated by the principles of geometry. Human knowledge about the makeup and functioning life on earth as much as the composition of the universes in galaxies and spaces advocate the important role of geometry. Besides humans, animals appear to be perfect practitioners of the geometric concepts. For an example, a preying tiger observes the speed of a target in motion and its path just to configure the optimal short chord and the right speed to run to capture the target. Ever since humans started to think about the natural world and the scope of geometry, they have been defining, refining, developing, and applying several geometric ideas and tools, the knowledge discovery process of geometry has been continuing and progressing. See Brunes (1967), Mandelbrot (1983) and Burton (2011) for an excellent narration of how geometry has been utilized by the humans in general since time immemorial. The medical and health professions are no exception but have much to gain and benefit by probing the applicability and advantages of the geometric concepts and their tools, especially of the triangles (see Posamentier, and Lehmann (2012) for details).

New probabilistic expressions are derived utilizing trigonometric relations among the corner points of a triangle. The conceptual contents of this article are versatile enough for different medical/health data analysis. This article emphasizes a vital role of triangle to understand and relate the events: never, once, and repeated occurrence of an illness. For an illustration, the occurrence of never, once, and repeated hay fever and asthma attack among twins of n_{MZF} =1032 monozygotic females (MZF), n_{MZM} =566 monozygotic males (MZM), n_{DZF} =595 di-zygotic females (DZF), and n_{DZM} =352 di-zygotic males (DZM) in a genetic study as reported in Hooper et al. (1990) are considered. In the end, suggestions are made for future research direction in this new geometric approach.

TRIGNOMETRIC VIEW OF DISCRETE DATA

Medical researchers and public health professionals often desire to learn from the collected data on an illness to comprehend its dynamics so that a successful treatment could be devised when it is feasible. The key to achieve it falls squarely on noticing both the apparent and hidden evidence in the data about the related uncertainty of the episodes of the illness. No analytic technique is more suiting than the geometric approach to comprehend the patterns. Hence, a geometric concept with related tools is worth developing and it is exactly done in this article.

Suppose that in a random sample $y_1, y_2, .., y_n$ of size *n* from a targeted population is collected about the incidences of an illness. Let $n_0 \geq 0$, $n_1 \geq 0$ and $n - n_0 - n_1 \geq 0$ among the sampled have experienced *never, once,* and *repeat* of the illness with unknown chances $0 \leq \pi_0 \leq 1$, $0 \leq \pi_1 \leq 1$ and $0 \leq \pi_2 = 1 - \pi_0 - \pi_1 \leq 1$ respectively. In a geometric format, these chances are vertical coordinates of a *triangle* (Figure 1 with $\pi_0 > \pi_1 > \pi_2$) with *vertices* A (**0,** π_0), B (1, π_1), and C (**2,** $1 - \pi_0 - \pi_1$) in a Cartesian graph as shown below. There are five of *six* different possible ways the triangles could be framed depending on whether $\pi_1 > \pi_2 > \pi_0$, $\pi_1 > \pi_0 > \pi_2$, $\pi_2 > \pi_1 > \pi_0$, $\pi_0 > \pi_2 > \pi_1$, or $\pi_2 > \pi_0 > \pi_1$.

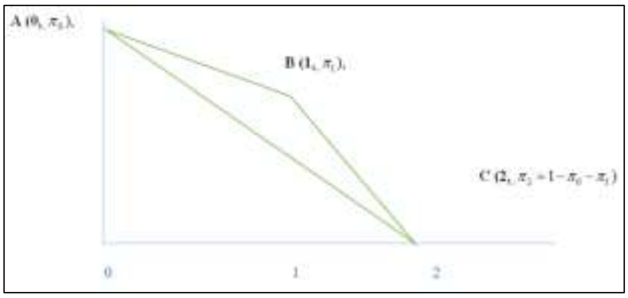

Figure 1: Triangle of never, once, and repeat events ($\pi_0 > \pi_1 > \pi_2$).

Incidentally, The adjective *Cartesian* glorifies the French mathematician and philosopher René Descartes (who used the name *Cartesius* in Latin in his first writing in 1637 about the locus of a point graphically). The idea of graphical system was independently developed in the same year by Pierre de Fermat, although Fermat worked in three dimensions but did not publish his discovery. Both authors used a single axis in their treatments with a variable length measured in reference to the axis.

First, we find the lengths: c, a, and b of AB, BC, and AC using the Euclidean distance $d = \sqrt{(m-p)^2 + (n-q)^2}$ between two points of the plane with Cartesian coordinates (m, n) and (p, q).

This article is first of its kind to investigate intrinsic geometric *relations* among the chances of *never, once,* and *repeat* of the illness and their practical implications in the knowledge discovery of the illness. In specific, intricacies and their implications of the vertices in the triangle are explored. Because more concern exists in a medical set-up about the repeated incidence, the vertex **C** (2, $\pi_2 = 1 - \pi_0 - \pi_1$) is given more priority over other two vertices in our discussions below, though similar arguments can be easily made for the other vertices and hence, are skipped in the article. Note that

$$a = \sqrt{1 + (1 - \pi_0 - 2\pi_1)^2}, \tag{1}$$

$$b = \sqrt{4 + (1 - 2\pi_0 - \pi_1)^2}, \tag{2}$$

and

$$c = \sqrt{1 + (\pi_1 - \pi_0)^2}. \tag{3}$$

Consequently, the cosine of the angles α, β, γ between the sides {AB and AC}, {AB and BC}, {BC and AC} are, after algebraic simplifications, become

$$\alpha = \cos^{-1}\left(\frac{c^2+b^2-a^2}{2cb}\right) = \cos^{-1}\left\{\frac{[(\pi_1-\pi_0)(1-2\pi_0)+2]}{\sqrt{[1+(\pi_1-\pi_0)^2][4+(1-2\pi_0-\pi_1)^2]}}\right\}, \quad (4.a)$$

$$\beta = \cos^{-1}\left(\frac{a^2+c^2-b^2}{2ac}\right) = \cos^{-1}\left\{\frac{[(\pi_1-\pi_0)(\pi_0+\pi_1-1)-1]}{\sqrt{[1+(1-\pi_0-2\pi_1)^2][1+(\pi_1-\pi_0)^2]}}\right\}, \quad (4.b)$$

$$\gamma = \cos^{-1}\left(\frac{a^2+b^2-c^2}{2ab}\right) = \cos^{-1}\left\{\frac{[4+(1-2\pi_0-\pi_1)^2+(1-\pi_0-2\pi_1)^2 - (\pi_1-\pi_0)^2]}{\sqrt{[1+(1-2\pi_0-\pi_1)^2][1+(1-\pi_0-2\pi_1)^2]}}\right\}. \quad (4.c)$$

Notice that the probabilities π_0 and π_1 determine all three angles. In some specific models, there might be intrinsic relation between π_0 and π_1 (which we will explore below for the binomial, Poisson, and inverse binomial as these are some popular models in medical/health studies).

Consequently, all three angles are determined solely by the probability of the null event itself (that is, $\pi_0 = \Pr[Y=0]$)). In addition, larger an angle refers the corresponding opposite side is longer meaning their vertices far distant. Individually or together, the other two angles will diminish suggesting and their opposite sides will be shorter (in which case, their vertices are in close proximity) and the events are almost equally likely. From the expression (4.a), we notice that the events $\Pr[Y=1]$ and $\Pr[Y=2]$ are almost equally likely if and only if

$$\Delta_{12} = (\pi_1-\pi_0)(2\pi_0-1)-2 = 0 \quad (5.a)$$

is satisfied, according to Pythagorean theorem. From the expression (4.b), it is clear that the events $\Pr[Y=0]$ and $\Pr[Y=2]$ are almost equally likely if and only if

$$\Delta_{02} = (\pi_1-\pi_0)(\pi_0+\pi_1-1)-1 = 0 \quad (5.b)$$

is satisfied. From the expression (4.c), it is evident that the events $\Pr[Y=1]$ and $\Pr[Y=0]$ are almost equally likely if and only if

$$\Delta_{01} = 4+(1-2\pi_0-\pi_1)^2+(1-\pi_0-2\pi_1)^2 - (\pi_1-\pi_0)^2 = 0 \quad (5.c)$$

is satisfied. Otherwise, it is simply a scalene triangle. If $c=a=b$, then it is *equilateral triangle* with their *angles* are equally 60^0. The *area* of the triangle ABC is

$$TriangleArea = \left|\frac{abSin\gamma}{2}\right| = \left|\frac{acSin\beta}{2}\right| = \left|\frac{bcSin\alpha}{2}\right|. \quad (6)$$

The *semi perimeter* of the triangle is

$$s = (c+a+b)/2. \quad (7)$$

The *diameter of the circumcircle* around the triangle is

$$\delta = \frac{c}{Sin\gamma} = \frac{b}{Sin\beta} = \frac{a}{Sin\alpha}. \quad (8)$$

The *radius* of the *inner circle* of the triangle ABC is

$$r = \left[\left|\sqrt{\frac{(c+b-a)(a+c-b)}{(a+b-c)(a+b+c)}}\right|\right]/2. \quad (9)$$

The incircle is the circle, which lies inside the triangle and touches all three sides. Its radius is called the *inradius*. The centers of the inner circle and circumcircle form an orthocentric system. An angle bisector of a triangle is a straight line through a vertex, which cuts the corresponding angle in half. By Cevian theorem in Indika (2011), the length of the median, angle bisector, altitude are respectively

$$median = 0.5\sqrt{\left|2(a^2+b^2)-c^2\right|}, \quad (10)$$

$$AngleBi\sec tor = [\sqrt{ab(a+b+c)(a+b-c)}]/(a+b), \quad (11)$$

and

$$Altitude = [\sqrt{\frac{(a+b+c)(c+b-a)}{(a+c-b)(a+b-c)}}]/2c \quad (12)$$

The name Cevian comes from the Italian engineer Giovanni Ceva, who proved a well-known theorem about cevians, which also bears his name. A median of a triangle is a straight line through a vertex and the midpoint of the opposite side, and divides the triangle into two equal areas. The three medians intersect in a single point, the triangle's centroid or geometric barycenter. The centroid of a rigid triangular object (cut out of a thin sheet of uniform density) is also its center of mass: the object can be balanced on its centroid in a uniform gravitational field. The centroid cuts every median in the ratio 2:1, i.e. the distance between a vertex and the centroid is twice the distance between the centroid and the midpoint of the opposite side. The midpoints of the three sides and the feet of the three altitudes all lie on a single circle, the triangle's nine-point circle. The remaining three points for which it is named are the midpoints of the portion of altitude between the vertices and the orthocenter. The radius of the nine-point circle is half that of the circumcircle. The above expressions are not quite interpretable in probability terms and hence, are not further pursued.

RESULTS FOR SPECIFIC DATA MODELS

In this section, the results of Section 2 are obtained for specific discrete probability models. The binomial, Poisson, and inverse binomial (geometric as a special case of the inverse binomial) are popularly utilized in medical/health studies.

FOR BINOMIAL MODEL

The total count, Y of a specific outcome in $n \geq 1$ independent and identical *Bernoulli cases* is called *binomial random variable* whose probability mass function is

$$\Pr[Y = y | n, \theta] = \frac{n!}{y!(n-y)!} \theta^y / (1+\theta)^n; \quad \theta > 0, n \geq 1, y = 0,1,2,...,n \quad (13)$$

where $\theta > 0$ denotes the *odds* of the specific outcome. Notice that $\pi_0 = 1/(1+\theta)^n$ and $\pi_1 = (\frac{1-\pi_0^{1/n}}{\pi_0^{1/n}})\pi_0$ which needs to be substituted in all expressions of section 2 for further simplifications for the binomial model. Only the *null event* $\pi_0 = \Pr[Y = 0 | n, \theta]$ controls all the properties of the *binomial triangle*. An illustrative example of genetic data is given in the next section for the binomial model.

FOR POISSON MODEL

The count, Y of a rare outcome is called *Poisson random variable* whose probability mass function is

$$\Pr[Y = y | \theta] = e^{-\theta} \theta^y / y!; \theta > 0, y = 0,1,2,...,\infty$$

where $\theta > 0$ denotes the *incidence rate*. Notice that $\pi_0 = e^{-\theta}$ and $\pi_1 = -\pi_0 \ln \pi_0$ which needs to be substituted in all expressions of section 2 for further simplifications for the Poisson model. Only the *null event* $\pi_0 = \Pr[Y = 0 | \theta]$ controls all the properties of the Poisson triangle.

FOR INVERSE BINOMIAL MODEL

The count, Y of the number of cases to be investigated until $r \geq 1$ cases with a specified outcome is called *inverse binomial random variable* whose probability mass function is

$$\Pr[Y = y | r, p] = \frac{\Gamma(r+y)}{y!\Gamma(r)} p^r (1-p)^y; \quad r \geq 1, 0 < p < 1, y = 0,1,2,...,\infty$$

where $0 < p < 1$ denotes the *chance of not observing* the specific outcome in a case. Notice that $\pi_0 = p^r$ and $\pi_1 = r[1 - e^{\frac{1}{r}\ln \pi_0}]\pi_0$ which needs to be substituted in all expressions of section 2 for further simplifications for the inverse binomial model. Only the *null event* $\pi_0 = \Pr[Y = 0 | r, p]$ controls all the properties of the *inverse binomial triangle*.

FOR GEOMETRIC MODEL

The *geometric* model is a special case of inverse binomial model with $r = 1$. That is, the count, Y of the number of cases to be investigated until just one case with a specified outcome is called *geometric random variable* whose probability mass function is

$$\Pr[Y = y | p] = p(1-p)^y; 0 < p < 1, y = 0,1,2,...,\infty$$

where $0 < p < 1$ denotes the *chance of not observing* the specific outcome in a case. Notice that $\pi_0 = p$ and $\pi_1 = [1 - \pi_0]\pi_0$ which needs to be substituted in all expressions of section 2 for further simplifications for the inverse binomial model. Only the *null event* $\pi_0 = \Pr[Y = 0 | p]$ controls all the properties of the *geometric triangle*.

ILLUSTRATIONS WITH GENETIC DATA FROM TWINS WITH RESPECT TO ASTHMA AND HAY FEVER

In this section, the results for binomial probability model are illustrated. In the illustration, a genetic data of twins in Hopper et al. [Genetic Epidemiology, 1990] about the occurrence of asthma and hay fever among four groups are used. The groups are: (1) monozygotic females (MZF), (2) monozygotic males (MZM), (3) di-zygotic females (DZF), and (4) di-zygotic males (DZM). In each group, the entries are random sample from a binomial probability mass function (13). The sample sizes are indicated by n_{MZF}, n_{MZM}, n_{DZF}, and n_{DZM} The entries of the binomial data for these four groups are displayed in Table 1. All four groups are of the same type: $\pi_0 > \pi_1 > \pi_2$.

Table 1: # twins with *never, once, and repeated* hay fever asthma attack among monozygotic females (MZF), monozygotic males (MZM), di-zygotic females (DZF), and di-zygotic males (DZM).

	Hay fever			prop	Angle		Hay fever			prop	Angle
Asthma	**0**	**1**	**2**			**Asthma**	**0**	**1**	**2**		
MZF						**MZM**					
0	385	258	137	0.75	α = 11°	0	304	100	59	0.81	α = 15°
1	41	70	74	0.17	β = 157°	1	22	20	22	0.11	β = 147°
2	12	9	46	0.06	γ = 12°	2	8	14	17	0.06	γ = 18°
prop	0.42	0.32	0.26			**prop**	0.59	0.23	0.18		
Angle	α = 1°	β = 178°	γ = 1°	n_{MZF} =1032		**Angle**	α = 8°	β = 164°	γ = 8°	n_{MZM} = 566	
DZF						**DZM**					
0	342	65	65	0.79	α = 11°	0	162	84	12	0.73	α = 7°
1	40	32	32	0.17	β = 156°	1	17	46	19	0.23	β = 164°
2	2	8	9	0.03	γ = 13°	2	3	6	3	0.03	γ = 9°
prop	0.65	0.18	0.17			**prop**	0.52	0.38	0.10		
Angle	α = 12°	β = 155°	γ = 13°	n_{DZF} =595		**Angle**	α = 4°	β = 171°	γ = 5°	n_{DZM} =352	

What is zygotic? A zygote is cell formed when two gamete cells are joined by means of sexual reproduction (see Klossner (2011) for details). In multicellular organisms, it is the earliest developmental stage of the embryo. In single-celled organisms, the zygote divides to produce offspring, usually through mitosis, the process of cell division. Zygotes are usually produced by a fertilization event between two haploid cells - an ovum (female gamete) and a sperm cell (male gamete) - which combine to form the single diploid cell. Such zygotes contain DNA derived from both parents, and this provides all the genetic information necessary to form a new individual.

According to (4.a, 4.b, and 4.c), the angles are calculated (see Table 1 for details) for all the four groups: MZF, MZM, DZF, and DZM. The angle $\beta = \angle ABC$ between sides AB and BC is 178^o, 164^o, 155^o, and 171^o in the case of *asthma incidence* respectively for the four groups: MZF, MZM, DZF, and DZM, and they are all *obtuse*. Likewise, the angle $\beta = \angle ABC$ between sides AB and BC is 157^o, 147^o, 156^o, and 164^o in the case of *hay fever* incidence respectively for the four groups: MZF, MZM, DZF, and DZM, and they are all *obtuse*. Consistently, in each group, the angle $\angle ABC$ is the largest meaning that the *Euclidean distance* BC between the *null event* $\Pr(Y = 0)$ and the *repeated event* $\Pr(Y = 2)$ is farthest in the binomial triangle of this genetic twin data with respect to *asthma* as well as *hay fever*. This further confirmed by the *length of the altitude* (12) from the vertex B to the side AC is the smallest in the case of *asthma incidence* as well as in the case of *hay fever incidence* for each group among MZF, MZM, DZF, and DZM, and they are 0.4. 0.5. 0.4, and 0.2 respectively for MZF, MZM, DZF, and DZM.

The *area, semi perimeter, median*, and *angle bisector* of the triangles are calculated using (6), (7), (10), and (11), and they are about the same for the groups MZF, MZM, DZF, and DZM. Hence, they are not further considered for discussion. The diameter of the circumcircle (8) and the radius of the inner circle (9) differ from each other among the groups MZF, MZM, DZF, and DZM but they are not interpretable in probability sense.

CONCLUSIONS

It is common in medical/health studies that the zero event is either structurally not observable or observed with inaccuracies. How should the geometric approach of data analysis be refined to suit either situation? Further research is needed in this direction. Also, different covariates might exist for each vertex of the triangle and how would a regression type analysis proceed to make use of the covariates to explain the Euclidean distances among the three vertices of the triangle. Such a methodology will be quite useful to medical researchers.

ACKNOWLEDGEMENTS

The author appreciates and thanks Texas State University as this research was performed with the support given to the author during his sabbatical leave in spring semester of 2015.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Brunes Tons. The Secrets of ancient geometry and its use. In: Brunes Tons, eds. A Book. Vol. 1 and 2. Copenhagen, Denmark: RHODOS Publication, Strandgade 36; 1967.
2. Burton David M. The history of mathematics: an introduction. In: Burton David M, eds. A Book. 7th ed. New York: McGraw-Hill; 2011.
3. Descartes René. Discourse on method, optics, geometry, and meteorology. In: Paul J. Oscamp, eds. A Book. Revised ed. Indianapolis, IN: Hackett Publishing; 2001.
4. Hopper JL, Hannah MC, Macaskill GT, Mathews JD. Twin concordance for a binary trait: III a bivariate analysis of hay fever and asthma. Genet Epidemiol. 1990;7:277-89.
5. Indika SA. A new theorem on any right-angled Cevian triangle. J World Feder Natl Maths Competit. 2011;24(02):29-37.
6. Klossner N. Jayne, Hatfield Nancy. Introduction. In: Klossner N. Jayne, Hatfield Nancy, eds. Introductory Maternity & Pediatric Nursing. 2nd ed. New York: Lippincott Williams & Wilkins Press; 2006: 107.
7. Mandelbrot Benoît B. The fractal geometry of nature. In: Mandelbrot Benoît B, eds. A Book. 5th ed. New York: Macmillan Press; 2012.
8. Posamentier Alfred S, Lehmann Ingmar. The secrets of triangles. In: Posamentier Alfred S, Lehmann Ingmar, eds. A Book. 4th ed. New York: Prometheus Books; 2012.

DOI: 10.18203/2320-6012.ijrms20150142
Cite this article as: Shanmugam R. Never, once, and repeated illness: a geometric view for insights and interpretations. Int J Res Med Sci 2015;3:1336-41.