

Index of unfairness

Ilija Barukčić¹

¹ Internist, Horandstrasse, DE-26441 Jever, Germany

Correspondence: Ilija Barukčić, Horandstrasse, DE-26441 Jever, Germany. Tel: 49-4466-333. E-mail: Barukcic@t-online.de

Received: January 12, 2019; Accepted: January 12, 2019; Published: January 12, 2019

Abstract

Objective: Objective scientific knowledge for many authors more valuable than true subjective belief is determined by research on primary data but a renewed analysis of already recorded or published data is common too. Ever since, an appropriate experimental or study design is an important and often a seriously underappreciated aspect of the informativeness and the scientific value of any (medical) study. The significance of study design for the reliability of the conclusions drawn and the ability to generalize the results from the sample investigated for the whole population cannot be underestimated. In contrast to an inappropriate statistical evaluation of a medical study, it is difficult to correct errors in study design after the study has been completed. Various mathematical aspects of study design are discussed in this article.

Methods: In assessing the significance of a fair study design of a medical study, important measures of publication bias are introduced. Methods of data or publication bias analysis in different types of studies are illustrated through examples with fictive data. Formal mathematical requirements of a fair study design which can and should be fulfilled carefully with regard to the planning or evaluation of medical research are developed.

Results.

Various especially mathematical aspects of a fair study design are discussed in this article in detail. Depending on the particular question being asked, mathematical methods are developed which allow us to recognize data which are self-contradictory and to exclude these data from systematic literature reviews and meta-analyses. As a result, different individual studies can be summed up and evaluated with a higher degree of certainty.

Conclusions

This article is intended to give the reader guidance in evaluating the design of studies in medical research even ex post which should enable the reader to categorize medical studies better and to assess their scientific quality more accurately.

Keywords: study design, quality, study, study type, measuring technique, publication bias

1. Introduction

Biostatistics or statistical analysis is based on the key idea that the observation of a sample of subjects which is drawn from a certain population can be used to arrive at meaningful conclusions or inferences about the population with high degree of accuracy. In biomedical research, the various aspects of clinical research and the credibility of the data from the study substantially depend on the study design (Grimes and Schulz, 2002) which is more important than analyzing its results. The study design should ensure that a null hypothesis is either rejected or accepted and the conclusions drawn reflect only the truth. In particular, a poorly analyzed study can be reanalyzed but a poorly designed study can recover only poorly. Especially a mal-designed study (inclusion and exclusion criteria and other factors) can have an impact on the quality of the study sample with the consequence that the same is not an appropriate representative of the population. Under such circumstances, other studies may fail to successfully replicate the results of this original study and the inferences drawn can be misleading and the statistical procedures used cannot help any more. The widespread and documented lack of completeness and transparency in the reporting of statistical methods used endangers the possibility that a new study carried out again can successfully reproduce sufficiently similar or the same results as the original study. In point of fact, more than half

(52%) of scientists surveyed believe that studies do not successfully reproduce sufficiently similar or the same results as the original studies (Baker, 2016). A careful re-evaluation of the statistical methods and other scientific means which underpin scientific inquiry and research goals appears to be necessary. While it is important to recognize the shortcoming of today's science, one issue which has shaped debates over studies published is the question: has the study measured what it set out to? Even if studies carried out can vary greatly in detail the data from the studies itself provide information about the credibility of the data.

2. Material and Methods

Systematic observation and experimentation, inductive and deductive reasoning are essential for any formation and testing of hypotheses and theories about the natural world. In one way or another, logically and mathematically sound scientific methods and concepts are crucial constituents of any scientific progress. When all goes well, different scientists at different times and places using the same scientific methodology should be able to generate the same scientific knowledge.

2.1 Definitions

2.1.1 Definition. The Sample Space

Let the sample space denote a set or a collection of all different possible outcomes of an experiment. Each possible single outcome x_t of the experiment is said to be a member of the sample space, or to belong to the space $S(X)$. A single outcome x_t of an experiment $S(X)$ is a member of $S(X)$ and denoted symbolically by the relation $x_t \in S(X)$. A set Y is contained in another set X if every element of the set Y also belongs to the set X . This relation is expressed symbolically by the expression $Y \subset X$, which is the set-theoretic expression for saying that Y is a subset of X . A subset of X that contains no elements is called an empty set, or null set, and it is denoted by the symbol \emptyset .

Example

The sample space of a six-sided die can be regarded as a set containing the six numbers 1, 2, 3, 4, 5, 6 each number representing a possible side of a die which occurs after a roll. Symbolically, the sample space $S(X)$ of a six-sided die can be written as

$$S(X) = \{1, 2, 3, 4, 5, 6\} \quad (1)$$

Let $\Psi(S(X))$ denote the wave function of the sample space $S(X)$. Let $\Psi^*(S(X))$ denote the complex conjugate of the wave function of the sample space $S(X)$. If nothing contrary is stated, it is

$$\Psi(S(X)) \times \Psi^*(S(X)) \equiv 1 \quad (2)$$

2.1.2 Definition. The Absolute Frequency

Let $f(x_t)$ denote the absolute frequency of the event x_t ($t = 1, \dots, n$), a subset of the sample space $S(X)$. Then

$$f(x_1) + f(x_2) + \dots + f(x_n) = S(X) \quad (3)$$

For our present purpose it is important to note that

$$\frac{f(x_1)}{S(X)} + \frac{f(x_2)}{S(X)} + \dots + \frac{f(x_n)}{S(X)} = \frac{S(X)}{S(X)} = 1 \quad (4)$$

2.1.3 Definition. The Relative Frequency

Let $f(x_t)$ denote a subset of the sample space $S(X)$. Let $p(X=x_t)$ denote the relative frequency or probability of an event x_t . Then

$$p(X = x_t) = \frac{f(x_t)}{S(X)} \quad (5)$$

Scholium.

Broadly speaking, there are numerous questions to be asked about the correct understanding of probability. Some authors will insist that a frequentist approach to the concept of probability is not intelligible enough. There are circumstances where the concept of *the degree of confidence* or other probability concepts are of use too. Be that as it may, for our purposes, the frequentist approach is very useful since the same enable us to relate facts and hypotheses of a particular kind with real word situations.

2.1.3 Definition. The Random Variables and Distributions

Let X denote a real-valued function defined on a sample space, a random variable, with a finite number of finite outcomes x_1 occurring with probability $p(X = x_1)$, x_2 occurring with probability $p(X = x_2)$, ..., x_n occurring with probability $p(X = x_n)$. The collection of all of these probabilities denotes the distribution of the discrete or continuous random variable X . A discrete distribution is characterized by its probability mass function (p. m. f.). A continuous distribution is characterized by its probability density function (p. d. f.). Let $E(x_t)$ denote the expectation value of a single event x_t . Let $E(x_t^2)$ denote the second moment expectation value of a single event x_t . In general, it is

$$\begin{aligned} E(x_t) &\equiv x_t \times p(X = x_t) \\ E(x_t)^2 &\equiv (x_t \times p(X = x_t))^2 \\ E(x_t^2) &\equiv x_t \times x_t \times p(X = x_t) \end{aligned} \quad (6)$$

Let $\sigma(x_t)^2$ denote the variance of a single event. Then

$$\begin{aligned} \sigma(x_t)^2 &\equiv E(x_t^2) - E(x_t)^2 \\ &\equiv (x_t \times x_t \times p(X = x_t)) - (x_t \times p(X = x_t))^2 \\ &\equiv (x_t \times x_t) \times (p(X = x_t) - (p(X = x_t))^2) \\ &\equiv (x_t \times x_t) \times p(X = x_t) \times (1 - p(X = x_t)) \end{aligned} \quad (7)$$

Let $E(X)$ denote the expectation value of the random variable X . It is then

$$\begin{aligned}
 E(X) &\equiv (x_1 \times p(X = x_1)) + (x_2 \times p(X = x_2)) + \dots + (x_n \times p(X = x_n)) \\
 &\equiv E(x_1) + E(x_2) + \dots + E(x_n) \\
 &\equiv \sum_{t=1}^{t=n} (x_t \times p(X = x_t)) \\
 &\equiv \Psi(X) \times X \times \Psi^*(X)
 \end{aligned} \tag{8}$$

and

$$\begin{aligned}
 E(X^2) &\equiv (x_1^2 \times p(X = x_1)) + (x_2^2 \times p(X = x_2)) + \dots + (x_n^2 \times p(X = x_n)) \\
 &\equiv E(x_1^2) + E(x_2^2) + \dots + E(x_n^2) \\
 &\equiv \sum_{t=1}^{t=n} (x_t^2 \times p(X = x_t)) \\
 &\equiv \Psi(X) \times X^2 \times \Psi^*(X)
 \end{aligned} \tag{9}$$

while $\Psi(X)$ denotes the wave function of the random variable X and $\Psi^*(X)$ denotes the complex conjugate of the wave function. Under conditions, where $X \times \Psi^*(X) = 1$ (Barukčić, 2016) it is

$$E(X) \equiv \Psi(X) \tag{10}$$

but not in general. Let $E(X^2)$ denote the expectation value of the second moment of a random variable X . Let $\sigma(x)^2$ denote the variance of the random variable X . Then

$$\sigma(X)^2 \equiv E(X^2) - E(X)^2 \tag{11}$$

2.1.4 Definition. Independence

Let A_t denote random variable at a Bernoulli trial (period of time) t . Let B_t denote another random variable at the same Bernoulli trial (period of time) t . Let $p(A_t)$ denote the probability of A_t . Let $p(B_t)$ denote the probability of B_t . Let $p(A_t \cap B_t)$ denote the joint probability of A_t and B_t . In the case of independence (de Moivre, 1718; Kolmogoroff, 1933) of A_t and B_t it is generally valid that

$$p(A_t \cap B_t) \equiv p(A_t) \times p(B_t) \tag{12}$$

2.1.5 Definition. The 2x2 Table

Let A_t denote random variable at a Bernoulli trial (period of time) t . Let B_t denote another random variable at the same Bernoulli trial (period of time) t . Let $p(A_t)$ denote the probability of A_t . Let $p(B_t)$ denote the probability of B_t . Let $p(A_t \cap B_t)$ denote the joint probability of A_t and B_t . In general it is $(a_t+b_t) = A_t$, $(c_t+d_t) = \underline{A}_t$, $(a_t+c_t) = B_t$, $(b_t+d_t) = \underline{B}_t$ and $a_t+b_t+c_t+d_t=N_t$. Equally, it is $B_t+\underline{B}_t = A_t + \underline{A}_t = N_t$. In this context, it is $p(a_t)=p(A_t \cap B_t)$ and in the same respect $p(A_t) = p(a_t)+p(b_t)$ or $p(A_t)= p(A_t \cap B_t)+ p(b_t) =p(A_t \cap B_t)+p(A_t \cap \underline{B}_t)$ while $p(A_t)$ is not defined as $p(a_t)$. In the same context, it is $p(B_t) = p(a_t)+p(c_t) = p(A_t \cap B_t) +p(c_t)$ and equally $p(\underline{B}_t) = 1- p(B_t) =p(b_t)+p(d_t)$. Since the joint probability of A_t and B_t is denoted in general by $p(A_t \cap B_t)$, it is $p(A_t \cap B_t) = p(A_t) - p(b_t)$ or $p(A_t \cap B_t)=p(B_t) - p(c_t)$ or in other words $p(B_t) + p(b_t) - p(c_t) = p(A_t)$. In general, it is $p(a_t)+p(c_t)+p(b_t)+p(d_t) = 1$ or equally $a_t + b_t + c_t + d_t = N_t$. The meaning of the abbreviations a_t , b_t , c_t , d_t , N_t are explained by following 2 by 2-table (Table 1).

Table 1. The sample space of a contingency table.

		Conditioned B_t		Total
		(Outcome)		
		Yes = 1	Not = +0	
Condition A_t (risk factor)	Yes =+1	a_t	b_t	A_t
	Not = +0	c_t	d_t	\underline{A}_t
Total		B_t	\underline{B}_t	N_t

2.1.5.a Definition. Index of unfairness

The index of unfairness (IOU) is defined as

$$IOU \equiv \left(\left(\frac{A_t + B_t}{N_t} \right) - 1 \right) \tag{13}$$

2.1.6 Definition. Sufficient Condition (*Conditio per Quam*)

The mathematical formula of the *sufficient* condition relationship (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) (*conditio per quam*) of a population is defined as

$$\begin{aligned}
 p(A_t \rightarrow B_t) &\equiv \frac{(a_t) + (c_t) + (d_t)}{N_t} = 1 \\
 &\equiv p(a_t) + p(c_t) + p(d_t) \\
 &\equiv p(B_t) + p(d_t) \\
 &\equiv p(a_t) + p(\underline{A}_t) \\
 &\equiv +1.
 \end{aligned} \tag{14}$$

and is used to prove the hypothesis: *if A_t then B_t* or is taken to express that *the occurrence of an event A_t is a sufficient condition* (Wertheimer, 1968; Gomes, 2009) *for existence or occurrence of an event B_t* . The occurrence of an event A_t is a sufficient condition for occurrence of the event B_t or B_t is a necessary condition for A_t . In other words, sufficient and necessary conditions (Wertheimer, 1968; Gomes, 2009) are converse relations.

2.1.7 Definition. The X^2 Test of Goodness of Fit of a Sufficient Condition

A random sample of observations can come from a particular distribution (sufficient condition distribution) but must not. The X^2 test of goodness-of-fit is an appropriate method for testing the null hypothesis that a random sample of observations comes from a specific distribution (i.e. the distribution of a sufficient condition) against the alternative hypothesis that the data have some other distribution. The additive property of X^2 distribution may sometimes be used as an additional test of significance. In this case, the continuity correction should be omitted from each X^2 value. Under conditions where the chi-square goodness of fit test cannot be used it is possible to use an approximate and conservative (one sided) confidence interval known as *the rule of three*. The X^2 distribution is a particular type of a gamma distribution and widely applied in the field of mathematical statistics. The applicability of using the Pearson chi-squared statistic in cases where the cell frequencies of a 2×2 contingency table are not greater than five is widely discussed (Fisher, 1922) in literature and the use of Yate's continuity correction (Yates, 1934) is proposed. However, studies provided evidence that incorporating Yate's continuity correction is not essential (Grizzle, 1967; Conover, 1974). Still, using *the continuity correction* (Yates, 1934), the chi-square value of a conditio per quam relationship is derived (Barukčić, 2018) as

$$X^2 \left((A_t \rightarrow B_t) | A_t \right) \equiv \frac{\left((b_t) - (1/2) \right)^2}{A_t} + 0 = 0 \quad (15)$$

or alternatively as

$$X^2 \left((A_t \rightarrow B_t) | B_t \right) \equiv \frac{\left((b_t) - (1/2) \right)^2}{B_t} + 0 = 0 \quad (16)$$

2.1.8 Definition. Necessary Condition (Conditio Sine Qua Non)

Among the many generally valid natural laws and principles under which nature or matter itself assures its own self-organization, a relationship between events denoted as a necessary (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) condition (a conditio sine qua non) is one among the most important. A necessary (or an essential) event or condition A_t for some event B_t is a condition that must be satisfied in order to obtain B_t . In this respect, to say that an event A_t with its own probability $p(A_t)$ is at the same (period of) time t a necessary condition for another event B_t with its own probability $p(B_t)$ is equivalent to say that it is impossible to have B_t without A_t . In other words, *without* A_t *no* B_t or the absence of A_t guarantees the absence of B_t . The mathematical formula of the *necessary* condition relationship (conditio sine qua non) of a population is defined as

$$\begin{aligned} p(A_t \leftarrow B_t) &\equiv \frac{(a_t) + (b_t) + (d_t)}{N_t} = 1 \\ &\equiv p(a_t) + p(b_t) + p(d_t) \\ &\equiv p(A_t) + p(d_t) \\ &\equiv p(a_t) + p(B_t) = p(a_t) + (1 - p(B_t)) \\ &\equiv +1. \end{aligned} \quad (17)$$

2.1.9 Definition. The X^2 Test of Goodness of Fit of a Necessary Condition

Under conditions where the chi-square goodness of fit test cannot be used it is possible to use an approximate and conservative (one sided) confidence interval known as *the rule of three*. Using *the continuity correction*, the chi-square value of a *conditio sine qua non* distribution (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) before changes to

$$X^2 \left((A_t \leftarrow B_t) | B_t \right) \equiv \frac{\left((c_t) - (1/2) \right)^2}{B_t} + 0 = 0 \quad (18)$$

Depending upon the study design, another method to calculate the chi-square value of a *conditio sine qua non* distribution (while using *the continuity correction*) is defined as

$$X^2 \left((A_t \leftarrow B_t) | \underline{A}_t \right) \equiv \frac{\left((c_t) - (1/2) \right)^2}{\underline{A}_t} + 0 = 0 \quad (19)$$

2.1.10 Definition. Exclusion (A_t Excludes B_t and Vice Versa Relationship)

The mathematical formula of the *exclusion* relationship (A_t excludes B_t and vice versa) of a population was defined (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) as

$$\begin{aligned} p(A_t | B_t) &\equiv \frac{(b_t) + (c_t) + (d_t)}{N_t} = 1 \\ &\equiv p(b_t) + p(c_t) + p(d_t) \\ &\equiv p(b_t) + p(\underline{A}_t) = p(b_t) + (1 - p(A_t)) \\ &\equiv p(c_t) + p(\underline{B}_t) = p(c_t) + (1 - p(B_t)) \\ &\equiv +1. \end{aligned} \quad (20)$$

and used to prove the hypothesis: A_t excludes B_t and vice versa. Why should A_t exclude B_t and vice versa? Under which conditions can such a relationship be given?

2.1.11 Definition. The X^2 Test of Goodness of Fit of the Exclusion Relationship

The chi square value with degree of freedom $2-1=1$ of the exclusion relationship (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) with a *continuity correction* can be calculated as

$$X^2 \left((A_t | B_t) | \underline{A}_t \right) \equiv \frac{\left((a_t) - (1/2) \right)^2}{\underline{A}_t} + 0 = 0 \quad (21)$$

Depending upon the study design, another method to calculate the chi-square value of a *conditio sine qua non* distribution is defined as

$$X^2 \left((A_t | B_t) | B_t \right) \equiv \frac{\left((a_t) - (1/2) \right)^2}{B_t} + 0 = 0 \quad (22)$$

The chi square Goodness of Fit Test of the exclusion relationship examines how well observed data compare with the expected theoretical distribution of an exclusion relationship.

2.1.12 Definition. The Mathematical Formula of the Causal Relationship k

The mathematical formula of the causal relationship k (Barukčić, 1989; Barukčić, 1997; Barukčić, 2005; Barukčić, 2006; Barukčić, 2009; Barukčić, 2011; Barukčić, 2012; Barukčić, 2016; Barukčić, 2017; Barukčić, 2018) is defined at every single event, at every single Bernoulli trial t, as

$$k(A_t, B_t) \equiv \frac{p(A_t \cap B_t) - (p(A_t) \times p(B_t))}{\sqrt{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} \quad (23)$$

where A_t denotes the cause and B_t denotes the effect. Under some certain circumstances, the chi-square distribution can be applied to determine the significance of causal relationship k. Pearson's concept of correlation is not identical with causation. Causation as such is not identical with correlation. This has been proved many times and is widely discussed in many publications.

2.1.13 Definition. The 95% Confidence Interval of the Causal Relationship k

A confidence interval (CI) of the causal relationship k calculated from the statistics of the observed data can help to estimate the true value of an unknown population parameter with a certain probability. Under some conditions, the 95% interval for the causal relationship k is derived as

$$\left\{ k(A_t, B_t) - \sqrt{\frac{5}{N}} ; k(A_t, B_t) + \sqrt{\frac{5}{N}} \right\} \quad (24)$$

2.1.14 Definition. The Binomial distribution

Historically, the binomial probability mass function of observing exactly x successes in n trials, with the probability of success on a single trial denoted by p and q = 1 - p is

$$p(X = x) = \left(\frac{n!}{x! \times (n - x)!} \right) (1 - q)^x \times (1 - p)^{n-x} \quad (25)$$

and was derived by the prominent Suisse mathematician Jacob Bernoulli (1655 - 1705) in his work *Ars Conjectandi* (Bernoulli, 1713). The mathematical formula to find probabilities in the binomial distributions may be very simple but to do the calculation itself can be pretty troublesome. A binomial distribution with parameters p and n = 1 is called the Bernoulli distribution with parameter p while x can take the values either +0 or +1. It is

$$p(X = x) = (1 - q)^x \times (1 - p)^{1-x} \quad (26)$$

Under conditions where $X = n$, the binomial distributions simplifies to

$$p(X = n) = \left(\frac{n!}{x! \times (n - x)!} \right) (1 - q)^x \times (1 - p)^{n-x} = \left(\frac{n!}{n! \times (n - n)!} \right) (1 - q)^n \times (1 - p)^{n-n} = (1 - q)^n \quad (27)$$

Under conditions where $X = 0$, the binomial distributions simplifies to

$$p(X = 0) = \left(\frac{n!}{x! \times (n-x)!} \right) (1-q)^x \times (1-p)^{n-x} = \left(\frac{n!}{0! \times (n-0)!} \right) (1-q)^0 \times (1-p)^{n-0} = (1-p)^n \quad (28)$$

Under certain circumstances, the Poisson distribution is an useful approximation to the binomial distribution with a very small success probability especially when the value of n is large and the value of p is close to 0. The binomial distribution can be approximated by the normal distribution too. The accuracy of such an approximation depends on several factors and first requires some pre-calculations. A rule of thumb using the normal distribution to approximate binomial probabilities is good if both $(N \times p) > 5-10$ and if $(N \times (1-p)) > 5-10$. Especially under conditions where the number of successes x is equal to the number of trials n or goes to n , such an approximation may provide inaccurate probabilities.

2.1.15 Definition. The Poisson distribution

The Poisson distribution, given previously by Abraham de Moivre (Moivre, 1733), is ascribed to Siméon Denis Poisson (1781–1840), a French mathematician, physicist, and engineer who published the same distribution 1837 in his work “*Recherches sur la probabilité des jugements en matière criminelle et en matière civile*” (Poisson, 1837). Ladislaus Bortkiewicz (Bortkiewicz, 1898) provided in 1898 one of the first practical applications of Poisson's distribution while investigating the number of soldiers in the Prussian army killed accidentally by horse kicks. A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$, if, for $x = 0, 1, 2, \dots$, the probability mass function of X is given by

$$p(X = x) = \left(\frac{\lambda^x}{x!} \right) \times e^{-\lambda} \quad (29)$$

where x is the number of times an event occurs in an interval and x can take values $0, 1, 2, \dots$, e is Euler's number (the number 2.71828..., the base of the natural logarithms) and $x!$ is the factorial of x or $x! = (x) \times (x-1) \times (x-2) \times \dots \times 2 \times 1$ and $\lambda = n \times p$ is the mean of the Poisson distribution. Many times, the Poisson distribution is applied to experimental conditions or situations with a large number of trials n while the occurrence of each event is very rare. Under conditions where $x=0$, it is

$$p(X = 0) = \left(\frac{\lambda^x}{x!} \right) \times e^{-\lambda} = \left(\frac{\lambda^0}{0!} \right) \times e^{-\lambda} = e^{-\lambda} \quad (30)$$

Scholium.

Suppose that the probability of a conditio sine qua non relationship $p(A_t \leftarrow B_t) \approx +1$. The probability the a conditio sine qua non relationship will not be given will be $\underline{p}(A_t \leftarrow B_t) = 1 - p(A_t \leftarrow B_t) \approx +0$. The expectation value that there is no conditio sine qua non relationship is $\underline{\lambda} = n \times \underline{p}(A_t \leftarrow B_t) = n \times (1 - p(A_t \leftarrow B_t))$. The probability follows as

$$p(X = 0) = p(X \leq 0) = \left(\frac{\lambda^x}{x!} \right) \times e^{-\lambda} = \left(\frac{\lambda^0}{0!} \right) \times e^{-\lambda} = e^{-\left(n \times (1 - p(A_t \leftarrow B_t)) \right)} \quad (31)$$

The probability of counting at least one rare Poisson event is 1 minus the probability of counting none, which is as

$$p(X \geq 1) = 1 - p(X = 0) = 1 - \left(\frac{\lambda^x}{x!} \right) \times e^{-\lambda} = 1 - \left(\frac{\lambda^0}{0!} \right) \times e^{-\lambda} = 1 - e^{-\left(n \times (1 - p(A_t \leftarrow B_t)) \right)} \quad (32)$$

2.1.16 Definition. Hypergeometric distribution

The hypergeometric distribution (Huygens and van Schooten, 1657; Pearson, 1899; Gonin, 1936) is defined by the parameters population size, event count in population, sample size and can be used to calculate the exact probability of an event even for small samples which are drawn from relatively small populations, without replacement. The hypergeometric distribution differs to some extent from the binomial distribution. In contrast to the hypergeometric distribution, the probability of a binomially distributed random variable is the same from trial to trial. While the chi square distribution is of limited value for samples drawn from relatively small populations, the hypergeometric distribution can be used to calculate the exact probabilities for samples drawn from relatively small populations and without replication and for large populations too. The probability of having randomly exactly a_t (Table 1) successes in N_t hypergeometric trials or the significance of the causal relationship k can be tested under conditions of sampling without replacement by the hypergeometric distribution too. The probability of having exactly a_t successes by chance in N_t hypergeometric experimental trials is given by

$$p(X = a_t) = \frac{\binom{A_t}{a_t} \times \binom{N_t - A_t}{B_t - a_t}}{\binom{N_t}{B_t}} \quad (33)$$

2.1.17 Definition. The Chebyshev inequality

Let X be a random variable with finite expected value $E(x)$ and finite non-zero variance $\sigma(x)^2$. Then for any real number $x > 0$, the probability $p(x)$ for each real number x calculated according to the Chebyshev's inequality (Bienaymé, 1846; Tchénychef, 1867) follows as

$$p(|X - \mu| \geq x) \leq \left(\frac{\sigma(X)^2}{x^2} \right) \quad (34)$$

The Chebyshev's inequality (also called the Bienaymé-Chebyshev inequality) provide us only very approximate values.

2.1.18 Definition. The rule of three

Under some specified conditions (i. e. the dataset analyzed is large enough or n , the sample size, is $n \sim 30$ and more), a Chi-square goodness of fit test (Pearson, 1900) is able to provide evidence whether a *sample* distribution observed is identical with a *theoretical* distribution expected. Formally, the Chi-square goodness of fit test is defined as $X^2 = ((\text{sample distribution}) - (\text{theoretical distribution}))^2 / (\text{theoretical distribution})$ or something like $X^2 = ((\text{observed}) - (\text{expected}))^2 / (\text{expected})$. An approximate and conservative (one sided) confidence interval as discussed by (Rumke, 1975; Louis, 1981; Hanley, 1983; Jovanovic and Levy, 1997) and known as *the rule of three* can be of practical value if the Chi-square goodness of fit test cannot be applied. The rule of three is known to be derived as

$$p_{\text{critical}} = 1 - \left(\frac{3}{N_t} \right) \quad (35)$$

and is one way to calculate the probability of events which occur with a probability near 1. Another and a very simple path to calculate the probability of an event can be performed by the following method.

2.1.19 Definition. The unknown population proportion π_{upper}

Tests of hypotheses concerning the sampling distribution of the sample proportion \mathbf{p} (i. e. *conditio sine qua non* p(SINE), *conditio per quam* p(IMP) et cetera) can be performed using the normal approximation. The calculation of the rejection region based on the sample proportion to construct a confidence interval for an unknown population proportion π_{upper} can be performed under conditions of sampling without replacement (Sachs, 1992) by the formula

$$p_{\text{critical upper}} = \left(p - \frac{1}{2 \times n} \right) - \left(Z \times \sqrt{\left(\frac{p \times (1-p)}{n} \right) \times \left(\frac{N-n}{N-1} \right)} \right) \quad (36)$$

while the term $((N-n)/(N-1))$ denotes *the finite population correction* (Isserlis, 1918).

2.1.20 Definition. Odds Ratio

The odds ratio (Fisher, 1935; Cornfield, 1951; Edwards, 1963; Mosteller, 1968), abbreviated as $OR(A_t, B_t)$, is a very commonly used measure of association for 2×2 contingency tables (Table 1) and given by

$$OR(A_t, B_t) \equiv \frac{a_t / b_t}{c_t / d_t} \equiv \frac{a_t \times d_t}{c_t \times b_t} \quad (37)$$

In addition, researchers are regularly relying on Odds ratio to gain some new knowledge. Still, we need to address some different aspect of Odds ratio itself to find out the straightforward contradictions and the deep theoretical inconsistency which is associated with Odds ratio. It turns out that we are ill-advised if we believe blindly, uncritically in Odds ratio.

Case $c_t = 0$ (conditio sine qua non relationship).

Many authors use the term “objective reality” to refer to something like nature or processes in nature which are independent of human mind and consciousness. Let's begin with a question: is the physical world inherently objective or is there some deeper reality beyond the same? A part of the real world which exists outside of ourselves is an interplay of various opposite dualities too and a real world doesn't hide anything if we can read the signs right. In particular, under similar circumstances, the same events, natural processes et cetera will be repeated. Thus far, to respond to some objections to the notion of “objective reality”, an event A_t can be a necessary condition of an event B_t , independently of our belief, independently of our mind and consciousness. Under conditions were $c_t=0$, there exists a *conditio sine qua non relationship* between A_t and B_t while in the same respect the Odds ratio collapses. To date, it is not generally accepted to divide by zero. The Odds ratio cannot speak about one of the natural, profound and far reaching relationships (i.e *conditio sine qua non*) but must pass over in silence on this relationship. Pagano & Gauvreau are quietly returning through the back door (Pagano & Gauvreau, 2018) to circumvent this fundamental problem of Odds ratio by adding 0.5 to the cells (Pagano & Gauvreau, 2018) a_t , b_t , c_t , d_t . This simple way to circumvent the inconsistency and spectacular methodological incompleteness of Odds ratio is fundamentally misleading. To date, a substantial amount of research data available is analyzed by the Odds ratio. The more serious difficulty of this point of view is that it appears to be impossible to rely on Odds ratio in principle.

Case $b_t = 0$.

Furthermore, under conditions were $b_t=0$, a *conditio per quam relationship* between A_t and B_t is given while the Odds ratio collapses again. For this reason, the Odds ratio is overshadowed by a deep theoretical inconsistency and appears not to be grounded on a seemingly sound piece of reasoning. More likely, the Odds ratio (OR) is nothing more but *Yule's coefficient of association* (Yule, 1900) $Q(A_t, B_t)$ re-written (Warrens, 2008) in a non-normalized form and given by

$$\begin{aligned} OR(A_t, B_t) &\equiv \frac{OR(A_t, B_t) - 1}{OR(A_t, B_t) + 1} \\ &\equiv \frac{\left(\frac{a_t \times d_t}{b_t \times c_t}\right) - 1}{\left(\frac{a_t \times d_t}{b_t \times c_t}\right) + 1} \\ &\equiv \frac{\left(\frac{(a_t \times d_t) - (b_t \times c_t)}{(b_t \times c_t)}\right)}{\left(\frac{(a_t \times d_t) + (b_t \times c_t)}{(b_t \times c_t)}\right)} \\ &\equiv \frac{\left((a_t \times d_t) - (b_t \times c_t)\right)}{\left((a_t \times d_t) + (b_t \times c_t)\right)} \end{aligned} \quad (38)$$

Under conditions where Yule's coefficient (Yule, 1900) of association $Q=0$, there is no association. Although severely and justifiably criticized especially by Karl Pearson (1857–1925), the long-time and rarely challenged

leader of statistical science and Heron (Pearson and Heron, 1913), Odds ratio is still regularly referred to. The standard error and 95% confidence interval of the Odds ratio (OR) can be calculated according to Altman (Altman, 1999). Given the severely limited character of odds ratio, the standard error of the log Odds ratio is calculated as

$$SE \left(\ln \left(OR(A_t, B_t) \right) \right) \equiv \sqrt{\left(\frac{1}{a_t} \right) + \left(\frac{1}{b_t} \right) + \left(\frac{1}{c_t} \right) + \left(\frac{1}{d_t} \right)} \quad (39)$$

where *ln* denotes the *logarithmus naturalis*. The 95% confidence interval of the odds ratio is given by

$$95 \% CI \equiv \exp \left(\ln \left(OR(A_t, B_t) \right) - \left(1.96 \times SE \left(\ln \left(OR(A_t, B_t) \right) \right) \right) \right)$$

to

$$\exp \left(\ln \left(OR(A_t, B_t) \right) + \left(1.96 \times SE \left(\ln \left(OR(A_t, B_t) \right) \right) \right) \right) \quad (40)$$

2.1.21 Definition. The Chi Square Distribution

The following critical values (Sachs, 1992) of the chi square distribution (Pearson, 1900) as visualized by Table 2 are used in this publication.

Table 2. The critical values of the chi square distribution (degrees of freedom: 1)

	p-Value	One sided X ²	Two sided X ²
	0.1000000000	1.642374415	2.705543454
	0.0500000000	2.705543454	3.841458821
	0.0400000000	3.06490172	4.217884588
	0.0300000000	3.537384596	4.709292247
	0.0200000000	4.217884588	5.411894431
	0.0100000000	5.411894431	6.634896601
The chi square distribution	0.0010000000	9.549535706	10.82756617
	0.0001000000	13.83108362	15.13670523
	0.0000100000	18.18929348	19.51142096
	0.0000010000	22.59504266	23.92812698
	0.0000001000	27.03311129	28.37398736
	0.0000000100	31.49455797	32.84125335
	0.0000000010	35.97368894	37.32489311
	0.0000000001	40.46665791	41.82145620

2.1.22 Definition. The p-value

A null hypothesis formulated before the performance of a scientific study should be either accepted or rejected. P-values are one of the useful statistical measures which enable us to some extent to compare the statistical plausibility and clinical relevance of the conclusions drawn about a study finding with respect to a random event. Historically, the evidence of the first use of the p-value in statistics dates back as far the late 17th century. The question of the p-value was addressed especially by John Arbuthnott (Arbuthnott, 1710) in 1710. Arbuthnot (1667 – 1735) examined birth records in London for each of the 82 years from 1629 to 1710 and compared the human sex ratio at birth to the null hypothesis of equal probability. About 100 years later, Pierre-Simon Laplace starts the Chapter V of his book “*Théorie analytique des probabilités*” (Laplace, 1812) with the computation of a p-value. In Chapter VI, of his book Laplace provided his famous study on the statistics of almost half a million births and demonstrated an excess of boys compared to girls. Laplace concluded by calculation of a p-value that the excess was a real effect. Formally, it was Karl Pearson who introduced the p-value (Pearson, 1900) as capital P. In point of fact, Fisher himself proposed in his influential book “*Statistical Methods for Research Workers*” (Fisher, 1925) the level p-value = 0.05 as a limit for statistical significance (Schervish, 1994). Many times, studies or experiments are investigating whether there is a difference between different experimental set-ups that the researchers are testing. In particular, a sample is drawn from a population, studied and the results are extrapolated to the population from where the sample was drawn. A condition or factor being studied can produces an effect or can make a difference but must not. In every experiment, the observed difference in the sample data must not reflect a true difference in the populations or in objective reality as such. To a certain extent, it is possible that a true null hypothesis is incorrectly rejected (type I error (or error of the first kind)). In other words, we falsely infer that something (i.e. H_0 ; *there is no difference*) is present when it actually it is not present. The probability of rejecting the null hypothesis given that the null hypothesis is true is called type I error rate or significance level, denoted by the Greek letter α (alpha). By convention, statisticians and journals suggest a significance level of $\alpha=5\%$ (Type I error) with the consequence (or potential consequence) that the difference observed is not due to chance but equally we have to accept to be fooled by randomness or subjective or objective random errors 1 time out of 20. In particular, the probability of incorrectly rejecting the null hypothesis or p (incorrectly rejecting the null hypothesis) = 5% is defined as being acceptable. A false null hypothesis should be rejected. Theoretically, it is possible to fail to reject a false null hypothesis (type II error or error of the second kind, β error). A false null hypothesis is rejected with the probability $1 - \beta$, denoted by the Greek letter β (beta). In an investigation, several statements based on the result of hypothesis tests are presented along with the associated p values. A hypothesis test should provide some help to decide whether the results of a study, based on a small sample, provide enough evidence against a claimed null hypothesis (denoted by H_0), with the consequence that it is reasonable to believe that in a larger target population, H_0 is false too. The strength of our evidence against H_0 is measured by the p-value. Still, there are some misunderstandings associated with the interpretation of a p value. In particular, a very small p value does provide strong evidence that H_0 is not true. In contrast to this, even as large p value does not provide real evidence that H_0 is true. Thus far and depending on the point of view, the p-value (Panagiotakos, 2008) is defined as the probability of obtaining a result equal to or more extreme than an actually observed result under the condition that a null hypothesis is valid. In general, it is

$$p\left((X \leq x)|H_0\right) + p\left((X > x)|H_0\right) \equiv 1 \quad (41)$$

The (left tailed) null and alternative hypothesis is under some circumstances as follows:

$$H_0 : p \geq p_0 \quad (42)$$

$$H_A : p < p_0$$

A left tailed p-value which is greater than or equal to α (p-value $\geq \alpha$) provides some evidence to *accept the null hypothesis* while a p-value calculated which is less than α (p-value $< \alpha$), support the decision to *reject the null hypothesis*. In other words, under the condition of the validity of the null-hypothesis, the left tailed p value can be calculated using the formula

$$p - value_{left\ tailed} = p\left((X \leq x)|H_0\right) \equiv 1 - p\left((X > x)|H_0\right) \quad (43)$$

Following Fisher's approach, a null hypothesis is never proved, but is possibly disproved. The testing of hypothesis is one of the means of examining the discrepancy between a null hypothesis and experimental data. In this context, a p-value of 0.10, means that a difference observed can be attributed to chance by 10%. To date by convention, the threshold of significance (i.e., α) is commonly set to 0.05, 0.01, 0.005, or 0.001. If the left tailed *p-value calculated is greater than or equal to α* (p-value $\geq \alpha$),

$$p - value_{left\ tailed} = p \left((X \leq x) | H_0 \right) \equiv 1 - p \left((X > x) | H_0 \right) \geq \alpha ? \quad (44)$$

then there is evidence to *accept the null hypothesis*. If the *p-value calculated is less than α* (p-value $< \alpha$), then there is evidence to *reject the null hypothesis*. It follows that

$$p \left((X > x) | H_0 \right) = 1 - (p - value_{left\ tailed}) = 1 - p \left((X \leq x) | H_0 \right) \quad (45)$$

The calculation of the probability $1 - p(X > x | H_0)$ enable us to calculate the left tail (lower) p-value. The (right tailed) null and alternative hypothesis is sometimes as follows:

$$H_0 : p \leq p_0 \quad (46)$$

$$H_A : p > p_0$$

A right tailed *p-value which is greater than or equal to α* (p-value $\geq \alpha$) provides some evidence to *accept the null hypothesis* while a *p-value calculated which is less than α* (p-value $< \alpha$), support the decision to *reject the null hypothesis*. The p-value tells the investigator the probability of obtaining a result *equal to or more extreme* than what was actually observed. In general, the *p-value for a right tail (upper) event* is given by

$$p - value_{right\ tailed} = p \left((X \geq x) | H_0 \right) \equiv 1 - p \left((X < x) | H_0 \right) \quad (47)$$

If the right tailed *p-value calculated is greater than or equal to α* (p-value $\geq \alpha$),

$$p - value_{right\ tailed} = p \left((X \geq x) | H_0 \right) \equiv 1 - p \left((X < x) | H_0 \right) \geq \alpha ? \quad (48)$$

then there is evidence to *accept the null hypothesis*. If the *p-value calculated is less than α* (p-value $< \alpha$), then there is evidence to *reject the null hypothesis*.

Example. Conditio sine qua non.

Suppose $x = 395$ as the number of times the *conditio sine qua non* relationship occurred in $n = 400$ trials. This random variable has the binomial distribution where p is the population parameter corresponding to the probability of success on any trial. The binomial distribution is used when there are exactly two *mutually exclusive outcomes* of a trial. The formula for the binomial probability mass function of observing exactly x successes in n trials, with the probability of success on a single trial denoted by p is

$$p(X = x) = \left(\frac{n!}{x! \times (n - x)!} \right) (p)^x \times (1 - p)^{n-x} \quad (49)$$

The probability of exactly $x=395$ events out of $n=400$ trials is $p(X = 395) = 0.0000412947$. The probability of exactly *not* $x=395$ events out of $n=400$ trials is $p(X <> 395) = 1 - p(X = 395) = 0.9999587053$. The probability of exactly, or more than $x=395$ events out of $n=400$ trials is calculated as $p(X \geq 395) = 0.0000533965$. The probability of less than $x=395$ events out of $n=400$ trials is calculated as $p(X < 395) = 0.9999466035$. The probability of more than $x=395$ events out of $n=400$ trials is calculated as $p(X > 395) = 0.0000121017$. Setting $\alpha = 0.05$, we have the cumulative probability of $x=395$ out of 400 events as

$$p(X \leq 395 | H_0) = p(X = 0) + p(X = 1) + \dots + p(X = 395) \quad (50)$$

The probability of exactly, or fewer than, $x=395$ events out of $n=400$ trials is $p = 0.9999878983$ or in other words

$$p(X \leq x | H_0) = \sum_{t=0}^{t=x} \left(\frac{n!}{t! \times (n - t)!} \right) (p)^t \times (1 - p)^{n-t} = \sum_{t=0}^{t=395} \left(\frac{n!}{t! \times (n - t)!} \right) (0.95)^t \times (1 - 0.95)^{n-t} = 0.9999878983 \quad (51)$$

The probability $p(X > 395) = 0.0000121017$ follows as

$$p(X > x | H_0) = 1 - \sum_{t=0}^{t=x} \left(\frac{n!}{t! \times (n - t)!} \right) (p)^t \times (1 - p)^{n-t} = 1 - \sum_{t=0}^{t=395} \left(\frac{n!}{t! \times (n - t)!} \right) (0.95)^t \times (1 - 0.95)^{n-t} = 0.0000121017 \quad (52)$$

Again, if $p \text{ value} \geq \alpha$ then accept H_0 ; if $p \text{ value} < \alpha$ then reject H_0 . We used a one-tailed test with null and alternative hypothesis and conclude with 95% confidence to accept the null hypothesis H_0 and reject the alternative hypothesis H_A since the $p \text{ value} = 0.9999878$ alternative hypothesis, because the associated p -value is less or equal to the level of significance α . In this context, the $p \text{ value}$ is thus the smallest level of significance to which the null hypothesis can still be rejected. Under some certain circumstances $n \times p \times (1 - p) > 9$. Another rule of thumb demands that the sample size n is “sufficiently large” and the binomial distribution can be approximated by the normal distribution if at least $n \times p \geq 5$ and if $n \times (1 - p) \geq 5$. If these conditions are met, then the binomial distribution can be treated as approximating the normal distribution and a z -test for significance can be performed.

Right tailed p value for a Binomial random variable according to Chebyshev inequality

Bernoulli's Theorem published in his Book *Ars Conjectandi* in 1713 (Bernoulli, 1713) and described by Poisson in 1837 (Poisson, 1837) under the name “*la loi des grands nombres*” (“The law of large numbers”) can be used to calculate the p value too. A Bernoulli distributed random variable x_t takes the value +1 with probability $p(x_t)$ and the value +0 with probability $q(x_t) = 1 - p(x_t)$ at each Bernoulli trial t . The Bernoulli distribution itself is a special case of the binomial distribution with $n = 1$. Let $X_n = (x_1 + x_2 + \dots + x_n)$ denote the sum of all independent, identically distributed (i.i.d.) random variables x_1, x_2, \dots, x_n . Let $H(X_n) = X_n/N_t$ denote the relative frequency of such a binomial distributed random variable with

$$E(H_n) = E\left(\frac{X_n}{N_t}\right) = \frac{E(X_n)}{N_t} = \frac{N_t \times p}{N_t} = p \tag{53}$$

while $E(H(X_n))$ denotes the expectation value of relative frequency. Let $\sigma(H(X_n))^2$ denote the variance of relative frequency as

$$\sigma(H_n)^2 = \left(\frac{1}{N_t^2}\right) \times \sigma(X_n)^2 = \sigma\left(\frac{X_n}{N_t}\right)^2 = \frac{p \times (1 - p)}{N_t} \tag{54}$$

Let X be a random variable with finite expected value $E(x)$ and finite non-zero variance $\sigma(x)^2$. In the case of *conditio sine qua non*, or *conditio per quam et cetera*, we define $E(x) = N$, $X = N \times (1 - \alpha)$ while α is the level of significance and $X_{\text{population}} = |X - E(x)|$. For any real number $x_{\text{observed}} > 0$, the p-value for each real number x can be calculated according to the Chebyshev's inequality. Our (right tailed) null and alternative hypotheses are as follows:

$$H_0 : |E(H_n) - 1| \geq \alpha \tag{55}$$

$$H_A : |E(H_n) - 1| < \alpha$$

The p-value according to the Chebyshev's inequality tell us again the probability of obtaining a result *equal to or more extreme* than what was actually observed. In general, the *p-value for a right tail (upper) event* is given by the Chebyshev's inequality as

$$p - \text{value}_{\text{right tailed}} = p\left(|E(H_n) - 1| \geq \alpha \mid H_0\right) \equiv 1 - p\left(|E(H_n) - 1| < \alpha \mid H_0\right) \leq \frac{\sigma(H_n)^2}{(\alpha)^2} \tag{56}$$

The p-value for the relative frequency according to Chebyshev's inequality (Bienaymé, 1846; Tchébychev, 1867) follows (Scheid, 1992) as

$$p(|E(H_n) - 1| \geq \alpha) \leq \left(\frac{\sigma(H_n)^2}{(\alpha)^2}\right) \leq \left(\frac{E(H_n) \times (1 - E(H_n))}{N_t \times (\alpha)^2}\right) \tag{57}$$

as long as $N_t \geq p \times (1 - p) / (\alpha^2)$. The Chebyshev's inequality (also called the Bienaymé-Chebyshev inequality) guarantees only an approximate p-values.

Example.

The sample size may be $N_t = 172$. Set the relative frequency as $E(H(X_n)) = p = 0,98255814$ and $\alpha = 0,05$. Since $p \times (1 - p) / (\alpha^2) = 6,855056787$ it is $N_t \geq p \times (1 - p) / (\alpha^2)$ and Chebyshev's inequality can be used. Within the population we expect the probability of 1. Our hypothesis is that the sample value does not deviate more than $\alpha = 0,05$ from the population. The p-value can be calculated as

$$p(|E(H_n) - 1| \geq \alpha) \leq \left(\frac{E(H_n) \times (1 - E(H_n))}{N_t \times (\alpha)^2} \right) \leq \left(\frac{0.98255814 \times (1 - 0.98255814)}{172 \times (0.05)^2} \right) = 0.03719038 \quad (58)$$

In general, if a p-value is *greater than or equal to* α (**p-value** $\geq \alpha$) then accept the null hypothesis. If the p-value calculated is less than α (**p-value** $< \alpha$), then reject the null hypothesis. Our p-value is less than α therefore we reject the null-hypothesis and accept the alternative hypothesis. The sample value deviates less than 5% from the population value.

Right tailed p value according to normal distribution

Our null-hypothesis is that the probability of the conditio sine qua non relationship is $\pi(A_t \leftarrow B_t)$. A random sample of the size n is drawn from the population. The absolute frequency of the conditio sine qua non relationship within the sample drawn was observed as $X(A_t \leftarrow B_t) = n \times p(A_t \leftarrow B_t)$ while $p(A_t \leftarrow B_t) = (X(A_t \leftarrow B_t) / n)$ denotes the relative frequency of the conditio sine qua non relationship within the sample. What is the probability that there will be $X(A_t \leftarrow B_t)$ or more cases of the conditio sine qua non relationship within the population? Obviously, the normalized variable z one sided right tailed (Sachs, 1992) becomes

$$z_{\text{one sided right tailed}} \equiv \frac{\left(p(A_t \leftarrow B_t) - \left(\frac{1}{(2 \times n)} \right) \right) - \left(\pi(A_t \leftarrow B_t) \right)}{\sqrt{\frac{\pi(A_t \leftarrow B_t) \times (1 - \pi(A_t \leftarrow B_t))}{n}}} \quad (59)$$

The continuity correction $(1/(2 \times n))$ becomes smaller as n becomes larger. This table 3 gives a probability that a statistic is less than Z (i.e. between negative infinity and Z) in other words

$$\Phi(z) \equiv \frac{1}{2 \times \pi} \int_{-\infty}^z e^{-t^2/2} dt \quad (60)$$

Table 3. Probability associated with a Z value.

Z	0	+0.01	+0.02	+0.03	+0.04	+0.05	+0.06	+0.07	+0.08	+0.09
1,00	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,10	0,86433	0,8665	0,86864	0,87076	0,87286	0,87493	0,87698	0,879	0,881	0,88298
1,20	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,30	0,9032	0,9049	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1,40	0,91924	0,92073	0,9222	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,50	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,60	0,9452	0,9463	0,94738	0,94845	0,9495	0,95053	0,95154	0,95254	0,95352	0,95449
1,70	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,9608	0,96164	0,96246	0,96327
1,80	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,90	0,97128	0,97193	0,97257	0,9732	0,97381	0,97441	0,975	0,97558	0,97615	0,9767
2,00	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,9803	0,98077	0,98124	0,98169
2,10	0,98214	0,98257	0,983	0,98341	0,98382	0,98422	0,98461	0,985	0,98537	0,98574
2,20	0,9861	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,9884	0,9887	0,98899
2,30	0,98928	0,98956	0,98983	0,9901	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,40	0,9918	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,50	0,99379	0,99396	0,99413	0,9943	0,99446	0,99461	0,99477	0,99492	0,99506	0,9952
2,60	0,99534	0,99547	0,9956	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,70	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,9972	0,99728	0,99736
2,80	0,99744	0,99752	0,9976	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,90	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
2,00	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,999

The following table 4 gives a probability that a statistic is greater than z.

$$f(z) = 1 - \Phi(z) \equiv 1 - \left(\frac{1}{2 \times \pi} \int_{-\infty}^z e^{-t^2/2} dt \right) \tag{61}$$

Table 4. Probability associated with a Z value.

z	+0	+0,01	+0,02	+0,03	+0,04	+0,05	+0,06	+0,07	+0,08	+0,09
1,00	0,15866	0,15625	0,15386	0,15151	0,14917	0,14686	0,14457	0,14231	0,14007	0,13786
1,10	0,13567	0,1335	0,13136	0,12924	0,12714	0,12507	0,12302	0,121	0,119	0,11702
1,20	0,11507	0,11314	0,11123	0,10935	0,10749	0,10565	0,10383	0,10204	0,10027	0,09853
1,30	0,0968	0,0951	0,09342	0,09176	0,09012	0,08851	0,08692	0,08534	0,08379	0,08226
1,40	0,08076	0,07927	0,0778	0,07636	0,07493	0,07353	0,07215	0,07078	0,06944	0,06811
1,50	0,06681	0,06552	0,06426	0,06301	0,06178	0,06057	0,05938	0,05821	0,05705	0,05592
1,60	0,0548	0,0537	0,05262	0,05155	0,0505	0,04947	0,04846	0,04746	0,04648	0,04551
1,70	0,04457	0,04363	0,04272	0,04182	0,04093	0,04006	0,0392	0,03836	0,03754	0,03673
1,80	0,03593	0,03515	0,03438	0,03362	0,03288	0,03216	0,03144	0,03074	0,03005	0,02938
1,90	0,02872	0,02807	0,02743	0,0268	0,02619	0,02559	0,025	0,02442	0,02385	0,0233
2,00	0,02275	0,02222	0,02169	0,02118	0,02068	0,02018	0,0197	0,01923	0,01876	0,01831
2,10	0,01786	0,01743	0,017	0,01659	0,01618	0,01578	0,01539	0,015	0,01463	0,01426
2,20	0,0139	0,01355	0,01321	0,01287	0,01255	0,01222	0,01191	0,0116	0,0113	0,01101
2,30	0,01072	0,01044	0,01017	0,0099	0,00964	0,00939	0,00914	0,00889	0,00866	0,00842
2,40	0,0082	0,00798	0,00776	0,00755	0,00734	0,00714	0,00695	0,00676	0,00657	0,00639
2,50	0,00621	0,00604	0,00587	0,0057	0,00554	0,00539	0,00523	0,00508	0,00494	0,0048
2,60	0,00466	0,00453	0,0044	0,00427	0,00415	0,00402	0,00391	0,00379	0,00368	0,00357
2,70	0,00347	0,00336	0,00326	0,00317	0,00307	0,00298	0,00289	0,0028	0,00272	0,00264
2,80	0,00256	0,00248	0,0024	0,00233	0,00226	0,00219	0,00212	0,00205	0,00199	0,00193
2,90	0,00187	0,00181	0,00175	0,00169	0,00164	0,00159	0,00154	0,00149	0,00144	0,00139
3,00	0,00135	0,00131	0,00126	0,00122	0,00118	0,00114	0,00111	0,00107	0,00104	0,001
3,10	0,00097	0,00094	0,0009	0,00087	0,00084	0,00082	0,00079	0,00076	0,00074	0,00071
3,20	0,00069	0,00066	0,00064	0,00062	0,0006	0,00058	0,00056	0,00054	0,00052	0,0005
3,20	0,00048	0,00047	0,00045	0,00043	0,00042	0,0004	0,00039	0,00038	0,00036	0,00035
3,40	0,00034	0,00032	0,00031	0,0003	0,00029	0,00028	0,00027	0,00026	0,00025	0,00024
3,50	0,00023	0,00022	0,00022	0,00021	0,0002	0,00019	0,00019	0,00018	0,00017	0,00017
3,60	0,00016	0,00015	0,00015	0,00014	0,00014	0,00013	0,00013	0,00012	0,00012	0,00011
3,70	0,00011	0,0001	0,0001	0,0001	0,00009	0,00009	0,00008	0,00008	0,00008	0,00008
3,80	0,00007	0,00007	0,00007	0,00006	0,00006	0,00006	0,00006	0,00005	0,00005	0,00005
3,90	0,00005	0,00005	0,00004	0,00004	0,00004	0,00004	0,00004	0,00004	0,00003	0,00003
4,00	0,00003	0,00003	0,00003	0,00003	0,00003	0,00003	0,00002	0,00002	0,00002	0,00002

Results

Theorem. The normalization of the variance of a single event

Mathematically, the probability that an event will occur is expressed as a number between +0 and +1 and can be defined in many different ways. For our purposes, the probability of event, which has a value or quantity x_t is represented by $p(x_t)$ and we define the probability that a single event has the value x_t at the Bernoulli trial t by the relationship

$$p(x_t) = \left(\frac{E(x_t)}{x_t} \right) = \left(\frac{x_t \times E(x_t)}{x_t \times x_t} \right) = \left(\frac{E(x_t^2)}{x_t^2} \right) \quad (62)$$

while $E(x_t)$ denotes the expectation value of a single event. Such a definition of probability assumes that every single event is associated with its own expectation value even under circumstances where $p(x_t) = 1$. Under these conditions it is equally $E(x_t) = x_t$. In other words, it is

$$E(x_t) = (x_t) \times p(x_t) = \left(\frac{E(x_t^2)}{x_t} \right) \quad (63)$$

or

$$E(x_t)^2 = (x_t^2) \times p(x_t)^2 = (x_t \times p(x_t))^2 \quad (64)$$

or

$$E(x_t^2) = (x_t^2) \times p(x_t) = (x_t) \times (x_t \times p(x_t)) = (x_t \times E(x_t)) \quad (65)$$

while the definitions above are independent of the distribution of x_t . The variance of a single event x_t denoted as $\sigma(x_t)^2$ is independent of the distribution of x_t and defined as

$$\sigma(x_t)^2 = E(x_t^2) - E(x_t)^2 = (x_t^2 \times p(x_t)) - (x_t \times p(x_t))^2 = x_t^2 \times (p(x_t) \times (1 - p(x_t))) \quad (66)$$

Claim.

In general, it is

$$\frac{E(x_t)^2}{E(x_t^2)} + \frac{\sigma(x_t)^2}{E(x_t^2)} = +1 \quad (67)$$

Proof.

$$\sigma(x_t)^2 \equiv E(x_t^2) - E(x_t)^2$$

Rearraing, it is

$$E(x_t)^2 + \sigma(x_t)^2 \equiv E(x_t^2)$$

Dividing, we obtain

$$\frac{E(x_t)^2}{E(x_t^2)} + \frac{\sigma(x_t)^2}{E(x_t^2)} \equiv \frac{E(x_t^2)}{E(x_t^2)} = +1 \quad (68)$$

or

$$\frac{E(x_t)^2}{E(x_t^2)} + \frac{\sigma(x_t)^2}{E(x_t^2)} \equiv +1$$

Quod erat demonstrandum.

Theorem. The exact probability of a single event I

Claim.

In general, it is

$$p(x_t) = 1 - \frac{\sigma(x_t)^2}{E(x_t^2)} \tag{69}$$

	Proof.	
$\sigma(x_t)^2$	\equiv	$E(x_t^2) - E(x_t)^2$
	<i>or</i>	
$\frac{E(x_t)^2}{E(x_t^2)} + \frac{\sigma(x_t)^2}{E(x_t^2)}$	$=$	1
	<i>or</i>	(70)
$p(x_t) = \frac{E(x_t)^2}{E(x_t^2)} = \frac{(x_t \times x_t \times p(x_t) \times p(x_t))}{(x_t \times x_t \times p(x_t))}$	\equiv	$1 - \frac{\sigma(x_t)^2}{E(x_t^2)}$

Quod erat demonstrandum.

Theorem. The exact probability of a single event II

Claim.

Under conditions of a Binomial distribution, it is

$$p(x_t) = 1 - \frac{\sigma(x)^2}{E(x)} \tag{71}$$

	Proof.	
$p(x_t) \times 1$	$=$	$p(x_t) \times 1$
$p(x_t)$	$=$	$p(x_t)$
	$=$	$0 + p(x_t)$
	$=$	$1 - 1 + p(x_t)$
	$=$	$1 - (1 - p(x_t))$
	$=$	$1 - \frac{p(x_t) \times (1 - p(x_t))}{p(x_t)}$
	$=$	$1 - \frac{N_t \times p(x_t) \times (1 - p(x_t))}{N_t \times p(x_t)}$
$p(x_t)$	\equiv	$1 - \frac{\sigma(x)^2}{E(x)}$

Quod erat demonstrandum.

Example. Chebyshev's inequality

According to Chebyshev's inequality, it is

$$p\left(|X - \mu| < \sqrt{E(X)}\right) \geq 1 - \frac{\sigma(x)^2}{E(X)} \quad (73)$$

while the number $E(X)$ is also called the mean of X or the expected value of X . The terms mean, expected value or expectation value are used interchangeably.

Theorem. The approximate probability p of a single event

Claim.

In general, the probability p of a single event is given approximately by

$$p = \sqrt[n]{e^{-(n \times (1-p))}} = \sqrt[n]{e^{-(n-X)}} = \sqrt[n]{e^{-\lambda}} \quad (74)$$

as long as the number of trials n goes to *positive infinity* ($n \rightarrow +\infty$) while $X = n \times p$ denotes the number of successes occurred anywhere among the n trials.

Proof.

In general, it is **+1 equal to +1** (*lex identitatis* (von Leibniz, 1765; Barukčić, 2016)) or

$$+1 = +1 \quad (75)$$

Multiplying by p , we obtain $1 \times p = 1 \times p$ or

$$p = p \quad (76)$$

where p denotes the probability of a single event. Let n denote something like the number of trials or the sample size et cetera. Performing the power operation, it is

$$(p)^n = (p)^n \quad (77)$$

According to mathematical requirements it is $p \equiv q$ or $p \equiv 1-p \equiv 1-q$ and $\lambda \equiv n \times p$ and $\underline{\lambda} \equiv n \times p \equiv n \times q \equiv n \times (1-p)$. Rearranging the equation before it is

$$(p)^n = (1-q)^n = \left(1 - \left(\frac{n \times q}{n}\right)\right)^n = \left(1 - \left(\frac{\lambda}{n}\right)\right)^n \quad (78)$$

Taking the limit as the number of trials as n goes to *positive infinity* ($n \rightarrow +\infty$), we obtain

$$\left(1 - \left(\frac{\lambda}{n}\right)\right)^n = \lim_{n \rightarrow +\infty} \left(\left(1 - \left(\frac{\lambda}{n}\right)\right)^n\right) \quad (79)$$

According to elementary (DeGroot et al., 2005) calculus it is

$$\lim_{n \rightarrow +\infty} \left(\left(1 - \left(\frac{\lambda}{n}\right)\right)^n\right) = e^{-\lambda} \quad (80)$$

as the number of trials n goes to *positive infinity* ($n \rightarrow +\infty$) the equation above simplifies as

$$(p)^n = e^{-\lambda} \quad (81)$$

In this context, the probability p in a sequence of n independent Bernoulli trials (experiments) with $q = 1 - p$ and $\underline{\lambda} = n \times (1-p)$ as the number of trials n goes to *positive infinity* ($n \rightarrow +\infty$) is given by

$$p = \sqrt[n]{e^{-(n \times (1-p))}} = \sqrt[n]{e^{-\lambda}} \quad (82)$$

Quod erat demonstrandum.

Example.

For a single trial, i.e., $n = 1$, it is

$$p = \sqrt[1]{e^{-(1 \times (1-p))}} = e^{-(1-p)} \quad (83)$$

and the probability is determined by its own other.

Example.

Suppose a team of Astronomers has investigated $n=10$ galaxies and found one black hole inside each galaxy, consequently it is $\lambda = 0$. The probability that every possible galaxy has a black hole can be calculated approximately as

$$p = \sqrt[n]{e^{-(n \times (1-p))}} = \sqrt[10]{e^0} = 1 \quad (84)$$

According to *the rule of three*, the probability that every galaxy does possess a black could be calculated as

$$p_{critical} = 1 - \left(\frac{3}{n}\right) = 1 - \left(\frac{3}{10}\right) = 0.7 \quad (85)$$

Example.

Suppose an investigation is performed with $n=100$ cases and controls. The probability of an event within the population is assumed to be $p=0.95$ and $\lambda = n \times q = n \times (1-p) = 100 \times (1-0.95) = 5$. What is the critical p ?

$$p_{critical} = \sqrt[n]{e^{-(n \times (1-p))}} = \sqrt[100]{e^{-5}} = 0.951229425 \quad (86)$$

The probability found within sample should not be lower than 0.951229425. Otherwise the data do not support the hypotheses that $p = 0.95$ or even more. Are such observations appropriate at all to justify some predictions about *observations we have not yet made or a reality, we are still not aware of* or may be even with regard to general claims which go far beyond the observed? The question is of course are we allowed to infer a hypothesis about the general situation based on the observation of such a limited sample? In other words, how (long) can we be uncertain about *the unknown*, the infinitely empty, the unobserved, on what ground and to what extent? One may object that any analysis of the notions of cause and effect is confronted by the unobserved and the not completely known too. On this view, how many galaxies are given within the universe known? We do not know for sure. How many of all galaxies do possess a black hole? We do not know for sure, either. Still, even such a small sample of observations justifies the conclusion and provides some degree of support but of course not the ultimate evidence for the truth that about 100 % of all galaxies possess a black hole. It is not the main goal of this paper to solve *the famous philosophical problem of induction* and inductive inference as introduces by *David Hume* in Book 1, part iii, section 6 in 1739 in his book "*A Treatise of Human Nature*" (Hume, 1739). However, in order to approach to the solution of this problem it is necessary to point out that under certain circumstances logic, mathematics and statistics are able to provide us to some extent with methods of direct inference even about the unknown.

Theorem.

In general, under conditions where $\mathbf{X} = \mathbf{x} = \mathbf{n}$, it is

$$p(X = n) = e^{-(n \times (1-p))} \quad (87)$$

were the probability of a single event is given by p and n is the sample size as the number of trials n goes to *positive infinity* ($n \rightarrow +\infty$).

Proof.

The binomial distribution is defined as

$$p(X = x) = \left(\frac{n!}{x! \times (n-x)!} \right) (1-q)^x \times (1-p)^{n-x} \quad (88)$$

were $p = 1 - q$. Under conditions where $\mathbf{X} = \mathbf{x} = \mathbf{n}$, it follows that

$$p(X = n) = \left(\frac{n!}{n! \times (n-n)!} \right) (1-q)^n \times (1-p)^{n-n} \quad (89)$$

or that

$$p(X = n) = (1-q)^n \quad (90)$$

Defining $\underline{\lambda} = n \times q = n \times (1-p)$, we obtain

$$p(X = n) = (1-q)^n = \left(1 - \frac{n \times q}{n} \right)^n = \left(1 - \left(\frac{\underline{\lambda}}{n} \right) \right)^n \quad (91)$$

Taking the limit as the number of trials as n goes to *positive infinity* ($n \rightarrow +\infty$), we obtain

$$\left(1 - \left(\frac{\underline{\lambda}}{n} \right) \right)^n = \lim_{n \rightarrow +\infty} \left(\left(1 - \left(\frac{\underline{\lambda}}{n} \right) \right)^n \right) \quad (92)$$

According to elementary (DeGroot et al., 2005) calculus it is

$$\lim_{n \rightarrow +\infty} \left(\left(1 - \left(\frac{\underline{\lambda}}{n} \right) \right)^n \right) = e^{-\underline{\lambda}} \quad (93)$$

In this context, the probability to obtain $x=n$ successes drawn with replacement from a population in a sequence of n independent Bernoulli trials (experiments) with $q = 1 - p$ and $\underline{\lambda} = n \times (1-p)$ as the number of trials n goes to *positive infinity* ($n \rightarrow +\infty$) is given by

$$p(X = n) = e^{-(n \times (1-p))} = e^{-\underline{\lambda}} \quad (94)$$

were the probability of a single event is given by p .

Quod erat demonstrandum.

As proved in this publication, the p -value for a *right tail (upper) event* under conditions were our expectation is that $\mathbf{X} = \mathbf{x} = \mathbf{n}$, is given by

$$p(X = n | H_0) \equiv 1 - p(X < n | H_0) = p(X \geq n | H_0) = e^{-(n \times (1-p))} \quad (95)$$

Theorem.

The *cumulative distribution function* abbreviated as $P(x)$ or as $F(x)$ or as $d.f.(x)$ or $c.d.f.(x)$ of every random variable X , regardless of whether the distribution of X is continuous, discrete or mixed, for each real number x is defined as

$$P(x) = F(x) = d.f.(x) = c.d.f.(x) = p(X \leq x) = p(X = x) + p(X < x) \quad (96)$$

for $-\infty < x < +\infty$.

Claim.

For every value x , it is

$$p(X > x) = 1 - p(X \leq x) \quad (97)$$

Proof.

For every value x , it is

$$p(X \leq x) + p(X > x) = 1 \quad (98)$$

and the theorem follows directly from the definition of the cumulative distribution function as

$$p(X > x) = 1 - p(X \leq x) \quad (99)$$

Quod erat demonstrandum.

Theorem.

The *p-value for a right tail (upper) event* is given by

$$p(X \geq x | H_0) \equiv 1 - p(X \leq x | H_0) + p(X = x | H_0) \quad (100)$$

Proof.

In general, it is

$$p(X \leq x | H_0) + p(X > x | H_0) \equiv 1 \quad (101)$$

or

$$p(X > x | H_0) \equiv 1 - p(X \leq x | H_0) \quad (102)$$

Adding $p(X=x|H_0)$, we obtain

$$p(X = x | H_0) + p(X > x | H_0) \equiv 1 - p(X \leq x | H_0) + p(X = x | H_0) \quad (103)$$

In general, it is $p(X \geq x | H_0) = p(X=x|H_0) + p(X>x|H_0)$. The *p-value for a right tail (upper) event* is given by

$$p(X \geq x | H_0) \equiv 1 - p(X \leq x | H_0) + p(X = x | H_0) = 1 - p(X < x | H_0) \quad (104)$$

Q. e. d.

Theorem.

The p-value for a right tail (upper) event under conditions where $\mathbf{X} = \mathbf{x} = \mathbf{n}$ is given by

$$1 - p((X < n)|H_0) = p((X = n)|H_0) \quad (105)$$

Proof.

In general, it is

$$p((X \leq x)|H_0) + p((X > x)|H_0) \equiv 1 \quad (106)$$

or

$$p((X = x)|H_0) + p((X < x)|H_0) + p((X > x)|H_0) \equiv 1 \quad (107)$$

Under conditions where $\mathbf{X} = \mathbf{x} = \mathbf{n}$, we obtain

$$p((X = n)|H_0) + p((X < n)|H_0) + p((X > n)|H_0) \equiv 1 \quad (108)$$

Mathematically, it is not possible that $X > n$. Thus far, $p((X > n)|H_0) = 0$. Under these assumptions, the equation before simplifies as

$$p((X = n)|H_0) + p((X < n)|H_0) \equiv 1 \quad (109)$$

Rearranging equation, it is

$$p((X = n)|H_0) \equiv 1 - p((X < n)|H_0) \quad (110)$$

Mathematically it is $p((X \geq x)|H_0) = 1 - p((X < x)|H_0)$. The p-value for a right tail (upper) event under conditions where $\mathbf{X} = \mathbf{x} = \mathbf{n}$, is given by

$$p((X = n)|H_0) \equiv 1 - p((X < n)|H_0) = p((X \geq n)|H_0) \quad (111)$$

Q. e. d.

The results one expects to obtain if some underlying assumption is true and the results observed while using some experimental data can differ by chance or systematically.

Theorem. *p value according to Poisson Distribution*

A binomial distribution is a sum of n independent Bernoulli random variables with the probability π . For very high or very low π , a binomial distribution is a very skewed distribution. Under conditions with very low π probability and very large n , the Poisson distribution may be used as an approximation to the binomial distribution. In practice it is possible not to observe a conditio sine qua non relationship within a sample even if within a population, such a relationship is given. Events like these can be accepted only under very limited circumstances and should be extremely small with the consequence that the law of rare events or Poisson limit theorem can be used to test the significance.

Claim.

The left tailed p-value of a Poisson distributed random variable (were $x = 0$) is given by

$$p(X \leq 0) = (1 - p(X > 0)) = p(X = 0) = e^{-\lambda} \quad (112)$$

Proof.

In general, it is

$$p(X \geq 0) = p(X = 0) + p(X > 0) = 1 \tag{113}$$

or

$$1 - p(X > 0) = p(X = 0) \tag{114}$$

Mathematically, the left tailed p-value is defined as $p(X \leq 0) = 1 - p(X > 0)$. Rearranging equation before, we obtain

$$p(X \leq 0) = (1 - p(X > 0)) = p(X = 0) \tag{115}$$

The Poisson distribution is given by

$$p(X = x) = \left(\frac{\lambda^x}{x!}\right) \times e^{-\lambda} \tag{116}$$

Under conditions where $x = 0$ we obtain

$$p(X = 0) = \left(\frac{\lambda^0}{0!}\right) \times e^{-\lambda} \tag{117}$$

and the left tailed p-value is given by

$$p(X \leq 0) = (1 - p(X > 0)) = p(X = 0) = e^{-\lambda} \tag{118}$$

Q. e. d.

Example

The (left tailed) null and alternative hypotheses may be as follows:

$$H_0 : p \geq 0.05 \quad (i. e. \text{ Conditio sine qua non: NO}) \tag{119}$$

$$H_A : p < 0.05 \quad (i. e. \text{ Conditio sine qua non: YES})$$

A left tailed p-value which is greater than or equal to α ($p\text{-value} \geq \alpha$) provides some evidence to *accept the null hypothesis* while a p-value calculated which is less than α ($p\text{-value} < \alpha$), support the decision to *reject the null hypothesis*.

The Poisson distribution is given by

$$p(X = x) = \left(\frac{\lambda^x}{x!}\right) \times e^{-\lambda} \tag{120}$$

were x is the number of times an event occurs in an interval and x can take values $0, 1, 2, \dots$, e is Euler's number (the number 2.71828..., the base of the natural logarithms) and $x!$ is the factorial of x or $x! = x \times (x-1) \times (x-2) \times \dots \times 2 \times 1$ and $\lambda = N \times p$ is a positive real number or the mean or equal to the expected number of rare occurrences of an event (i. e. no *conditio sine qua non* relationship observed). Under the condition of the validity of the null-hypothesis, the left tailed Poisson p value can be calculated using the formula

$$p - \text{value}_{\text{left tailed}} = p((X \leq x) | H_0) \equiv \sum_{t=0}^x \left(\frac{(N \times p)^t}{t!}\right) \times e^{-(N \times p)} \tag{121}$$

The sample size is again $N_t = 172$. The relative frequency of the *conditio sine qua non* relationship is $E(H(X_n)) = 169/172 = 0.98255814$ and $p = 0.05$. In other words, *the conditio sine qua non relationship was not*

observed in toto in 3 out of 172 cases. According to our left tailed hypothesis, we are of the opinion that this probability is greater or equal to 0.05. The left p-value can be calculated as

$$p - \text{value}_{\text{left tailed}} = p((X \leq x)|H_0) \equiv \sum_{t=0}^{x=3} \left(\frac{(N \times p)^t}{t!} \right) \times e^{-(N \times p)} = 0.02809258031854000000 \quad (122)$$

We must reject the null-hypothesis. *The probability that a conditio sine qua non relationship within the population will not be observed* is less than 0.05.

Theorem. Distribution and anti distribution

Suppose that S defines the *sample space* of an experiment completely. Let a real-valued function (*a random variable*) X which is defined on the sample space S assign a real number X(s) to each possible outcome $s \in S$ in a particular experiment. The distribution of the random variable of X is defined as the collection of all probabilities $p(X \in A)$ for all subsets A of the real numbers. A *discrete random variable* is defined as a random variable X which can take only a finite number of k different values x_1, \dots, x_k or at most, an infinite sequence of x_1, x_2, \dots . The distribution of a discrete random variable X is defined as the *probability mass function* and abbreviated as $p(x)$ or p. m. f.(x) of X, namely $p(x) = \text{p. m. f.}(x) = p(X = x)$ for all x in the set of possible values.

A random variable X which can take every value in an interval is called a *continuous random variable*. A continuous distribution is defined by its own *probability density function* (p.d.f.) of the distribution of X for every interval (a,b) as

$$p(a < X \leq b) \equiv \int_a^b f(x) dx \quad (123)$$

Continuous random variables satisfy the condition $p(X=x)=0$. In practical problems it may sometimes be necessary to consider a distribution as a mixture of a continuous distribution and a discrete distribution. Again, the *cumulative distribution function* abbreviated as P(x) or as F(x) or as d.f.(x) or c.d.f.(x) of every random variable X, regardless of whether the distribution of X is continuous, discrete or mixed, for each real number x is defined as

$$P(x) = F(x) = d.f.(x) = c.d.f.(x) = p(X \leq x) = p(X = x) + p(X < x) \quad (124)$$

for $-\infty < x < +\infty$.

Claim.

For every value x, the *anti distribution of x*, denoted as $p(X \neq x)$, is determined as

$$p(X \neq x) = 1 - p(X = x) = p(X < x) + p(X > x) \quad (125)$$

Proof.

For every value x, regardless of whether the distribution of X is continuous, discrete or mixed, it is

$$p(X \leq x) + p(X > x) = 1 \quad (126)$$

Since $p(X \leq x) = p(X=x) + p(X < x)$, the equation before can be rearranged as

$$p(X < x) + p(X = x) + p(X > x) = 1 \quad (127)$$

for $-\infty < x < +\infty$. Rearranging again, we obtain

$$p(X < x) + p(X > x) = 1 - p(X = x) \quad (128)$$

for $-\infty < x < +\infty$. We define the anti-distribution of x as $p(\underline{x}) \equiv p(X \neq x) \equiv p(X < x) + p(X > x)$ as the distribution for every value of *anti x* denoted as \underline{x} or as the *anti distribution of x* as

$$p(X \neq x) = p(X < x) + p(X > x) = 1 - p(X = x) \quad (129)$$

for $-\infty < x < +\infty$.

Quod erat demonstrandum.

Example.

The anti binomial distribution can be derived as

$$p(X \neq x) = 1 - p(X = x) = 1 - \left(\frac{n!}{x! \times (n-x)!} \right) (p)^x \times (1-p)^{n-x} \quad (130)$$

For n=25 and p=0.15 we obtain the following figure.

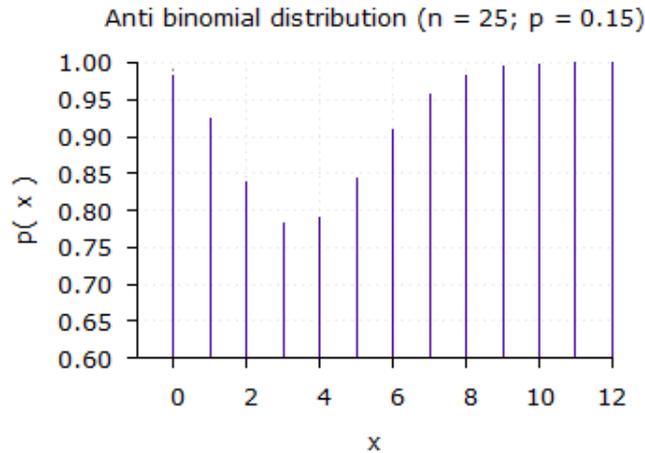


Figure 1. Anti-binomial distribution

The probability density of an *anti normal* (or *anti Gaussian* or *anti Gauss* or *anti Laplace–Gauss*) distribution follows as

$$p(X \neq x) = 1 - p(X = x) = 1 - \left(\left(\frac{1}{\sqrt{2 \times \pi \times \sigma(x)^2}} \right) \times \left(e^{-\left(\frac{x-\mu}{2 \times \sigma(x)^2} \right)} \right) \right) \quad (131)$$

where μ denotes the mean or expectation of the distribution and $\sigma(x)^2$ is the variance. For $\mu=1.0$ and $\sigma=1.5$ we obtain the following figure of an anti normal distribution.

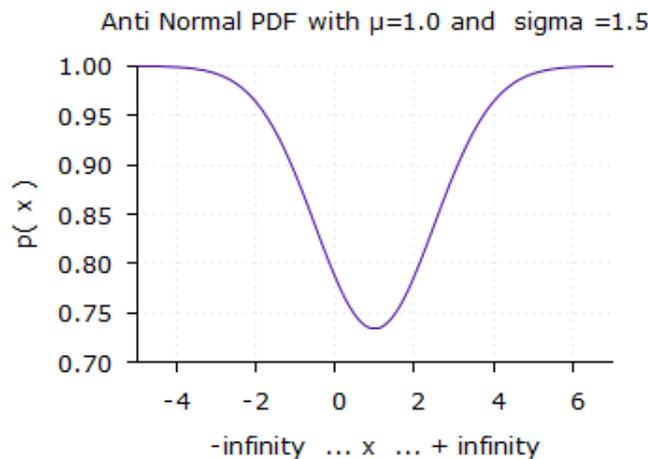


Figure 2. Anti normal distribution.

The (anti) normal distribution is useful because of the unofficial sovereign and the foundation of any statistics and probability theory, the *central limit theorem*. Any average of enough independent copies of a random variable will result nearly in a normal (Gaussian) distribution. The French-born mathematician Abraham de Moivre (1667 – 1754) while working on “Bernoulli’s Law of Large Numbers” (Moivre, 1718), the main theorem of Jakob

Bernoulli's *Ars conjectandi* (Bernoulli, 1713) published in the year 1733 the first historical pre-work on the central limit theorem (Moivre, 1733). After the monumental work “*Théorie analytique des probabilités*” (Laplace, 1812) of the famous French mathematician Pierre-Simon Laplace published in 1812 and the very precisely prove provided 1901 by the Russian mathematician Aleksandr Lyapunov (Lyapunov, 1901), the Hungarian born mathematician George Pólya coined in 1920 the German term “*zentraler Grenzwertsatz*” (Pólya, 1920) or *central limit theorem*. In a similar way, anti distributions of other distributions can be derived as demonstrated before.

The anti Poisson distribution

A discrete random variable X is said to have a Poisson distribution with parameter $\lambda > 0$, if, for $x = 0, 1, 2, \dots$, the probability mass function of X is given by

$$p(X = x) = \left(\frac{\lambda^x}{x!}\right) \times e^{-\lambda} \tag{132}$$

where x is the number of times an event occurs in an interval and x can take values 0, 1, 2, ..., e is Euler's number (the number 2.71828..., the base of the natural logarithms) and x! is the factorial of x or $x! = x \times (x-1) \times (x-2) \times \dots \times 2 \times 1$. Many times, the Poisson distribution is applied to experimental conditions or situations with a large number of trials n while the occurrence of each event is very rare. The probability density of an *anti Poisson distribution* is given by

$$p(X \neq x) = 1 - p(X = x) = 1 - \left(\left(\frac{\lambda^x}{x!}\right) \times e^{-\lambda}\right) \tag{133}$$

where x is the number of times a very rare event was observed. Thus far, let the probability p of an event in n Bernoulli trials be extremely near 1 or it is $p \sim 1$. In the same context, the probability, denoted by q, that the same event *will not occur* in n Bernoulli is extremely small and will be very near to zero or $q \sim 0$. It is $p + q = 1$ and $q = 1 - p$ and the expectation value of the very rare events is $\underline{\lambda} = n \times q = n \times (1 - p)$. In an experiment we observed $X = n \times p$ events. The rate of very rare events which should not have occurred is $x = n - X$ or in detail $x = n - X = n - (n \times p) = n \times (1 - p) = \underline{\lambda}$. In other words, if we know that the probability of x very rare $p(X=x)$, we know equally that the probability $p(X \neq x)$ of very often events/non-events is $p(X \neq x) = 1 - p(X=x)$. For $\mu = 2.0$ we obtain the following figure of an anti Poisson distribution.

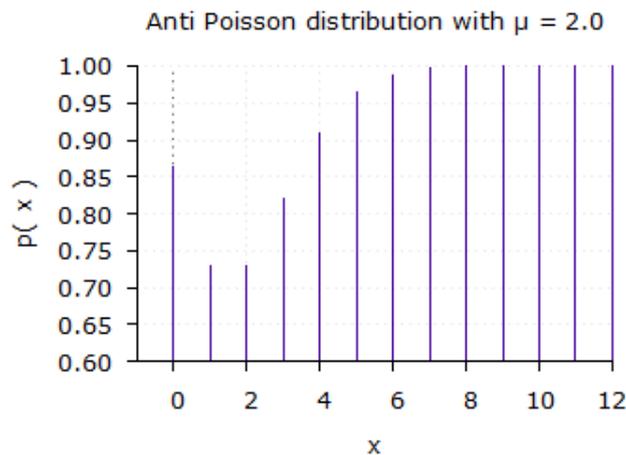


Figure 3. Anti Poisson distribution.

Consequently, the anti Poisson distribution above *under experimental conditions were the expectation value λ is equal to the number of rare events x*, i. e. where $(\lambda = x) > 0$ simplifies as

$$p(X \neq x) = 1 - p(X = x) = 1 - \left(\left(\frac{x^x}{x!}\right) \times e^{-x}\right) \tag{134}$$

where x indicates the expected (or average) number of occurrences of a very rare event. This very simplified form of the Poisson distribution can be called *the distribution of the likely*. Under conditions where $\lambda=n$ the anti Poisson distribution simplifies as

$$p(X \neq x) = 1 - p(X = x) = 1 - \left(\left(\frac{n^x}{x!} \right) \times e^{-n} \right) \quad (135)$$

and were ($\lambda=x=n$) as

$$p(X \neq n) = 1 - p(X = n) = 1 - \left(\left(\frac{n^n}{n!} \right) \times e^{-n} \right) \quad (136)$$

Example. The distribution of likely events

Suppose that, on the average, 1999 houses in 2000 in a certain district *are free of fire* or not burning during a year. If $n = 4000$ houses are in that district, what is the probability that exactly 3995 houses will stay *free of fire* or will not have a fire during the year. We focus on the fact that 1999 houses from 2000 houses *will not burn*, which is not a Poisson distributed random variable. In turn, it is insightful to point to the fact that 1 out of 2000 houses will have a fire which is a very rare event and Poisson distributed. As is so often the case, it is a matter of personal taste whether a glass is treated as half full or whether the same glass is treated as half empty. The anti Poisson distribution can be used to calculate the probability. Since 1999 houses have no fire, we know that 1 house in 2000 has fire or it is $q = 1-p = (1/2000)$ or $\lambda = n \times (1-p) = n \times q = 4000 \times (1/2000) = 2$. The probability that exactly 3995 houses will *have no fire* during a year means that exactly 5 houses or $4000-3995 = 5$ houses will have a fire. In other words, the probability that exactly 3995 houses in 4000 will *have no fire* during a year is

$$p(X \neq x) = 1 - \left(\left(\frac{\lambda^x}{x!} \right) \times e^{-\lambda} \right) = 1 - \left(\left(\frac{2^5}{5!} \right) \times e^{-2} \right) = 1 - 0.036089408863097 = 0.963910591 \quad (137)$$

and extremely near 1 and equivalent with the rare event 1 minus the probability that exactly 5 houses in 4000 houses will have a fire ($p=0.036089408863097$). Ultimately, under conditions were an event occurs its own complementary event does not occur or it is $p + q = 1$, the two terms are more or less interchangeable and it remains a matter of personal taste what is understood as p and what is taken as q .

Theorem. The Chi-square goodness-of fit test of a necessary condition

Unfortunately, there is always the possibility that the results of a study may be wrong and sometimes, a difference observed during an investigation is just the result of random subjective or objective errors or random effects. A statistical test is more or less about managing such and similar risks by the tools of probability theory and not about certainty. In point of fact, a true null hypothesis (there is no difference) should be accepted. Thus far we assume that a null hypothesis (H_0) is true.

Let $p(A_t)$ denote the probability of a condition (i.e. a risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability of A_t and B_t . The relationship between A_t and B_t is determined in many ways, both can be independent of each other too. Still, under conditions were the relationship between an event A_t and another event B_t is determined by a **necessary condition**, a *conditio sine qua non*, it is $p(A_t \leftarrow B_t) = p(A_t \text{ and } B_t) + (1 - p(B_t)) = p(d_t) + p(A_t) = 1$ and equally **$p(A_t \text{ and } B_t) = p(B_t)$** .

Claim.

In general, a mathematical formula of the Chi-square goodness-of fit test of a *necessary* condition can be derived as

$$X^2 \left((A_t \leftarrow B_t) | B_t \right) = \frac{(-c_t)^2}{B_t} = 0 \tag{138}$$

Proof.

The *conditio sine qua non* relationship of a population is defined as

$$p(A_t \leftarrow B_t) \equiv p(a_t) + p(b_t) + p(d_t) = (p(a_t) + p(b_t)) + p(d_t) = p(a_t) + (p(d_t) + p(b_t)) = 1 \tag{139}$$

(Table 1) or as

$$p(A_t \leftarrow B_t) \equiv p(a_t) + p(b_t) + p(d_t) = p(A_t) + p(d_t) = p(a_t) + (\underline{B}_t) = 1 \tag{140}$$

To see how this applies to the theorem above, let's simplify the equation before as

$$p(a_t) = 1 - (p(b_t) + p(d_t)) = 1 - (p(\underline{B}_t)) \tag{141}$$

or as

$$p(a_t) = 1 - (p(\underline{B}_t)) \tag{142}$$

In general it is $p(B_t) = 1 - p(\underline{B}_t)$. Ultimately, for this reason, a *conditio sine qua non* relationship simplifies under the press of mathematics as

$$p(a_t) = p(B_t) \tag{143}$$

Multiplying by N_t , we obtain

$$N_t \times p(a_t) = N_t \times p(B_t) \tag{144}$$

which is equivalent with

$$a_t = B_t \tag{145}$$

Rearranging, it is

$$a_t - B_t = 0 \tag{146}$$

or

$$(a_t - B_t)^2 = 0^2 \tag{147}$$

Dividing by B_t , it is

$$\frac{(a_t - B_t)^2}{B_t} = \frac{0}{B_t} = 0 \quad (148)$$

or

$$\frac{(a_t - B_t)^2}{B_t} + 0 = \frac{0}{B_t} + 0 = 0 \quad (149)$$

More precisely, it is

$$\frac{(a_t - B_t)^2}{B_t} + \frac{\left((b_t + d_t) - (b_t + d_t) \right)^2}{(b_t + d_t)} = \frac{0}{B_t} + 0 = 0 \quad (150)$$

or

$$X^2 \left((A_t \leftarrow B_t) | B_t \right) = \frac{(a_t - B_t)^2}{B_t} = \frac{0^2}{B_t} = 0 \quad (151)$$

It is $\mathbf{B}_t = \mathbf{a}_t + \mathbf{c}_t$. More broadly, the equation reduces to

$$X^2 \left((A_t \leftarrow B_t) | B_t \right) = \frac{\left(a_t - (a_t + c_t) \right)^2}{B_t} = 0 \quad (152)$$

Finally, the clearness, beauty and simplification provided by equation before yields the Chi-square goodness of fit test of a necessary condition *without using the continuity correction*, as

$$X^2 \left((A_t \leftarrow B_t) | B_t \right) = \frac{(-c_t)^2}{B_t} = 0 \quad (153)$$

Quod erat demonstrandum.

Depending upon personal taste, another method to calculate the chi-square value of a *conditio sine qua non* relationship *with the continuity correction* as demonstrated before can be derived as

$$X^2 \left((A_t \leftarrow B_t) | \underline{A}_t \right) \equiv \frac{\left((c_t) - (1/2) \right)^2}{\underline{A}_t} + 0 = 0 \quad (154)$$

Theorem. The Chi-square goodness-of fit test of a sufficient condition

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability of A_t and B_t . Under conditions where the relationship between A_t and B_t is determined by a **sufficient condition** it is $p(A_t \rightarrow B_t) = p(A_t \text{ and } B_t) + (1 - p(A_t)) = 1$ and it is equally $p(A_t \text{ and } B_t) = p(A_t)$.

Claim.

In general, a mathematical formula of the Chi-square goodness-of fit test of a *sufficient condition* can be derived as

$$X^2 \left((A_t \rightarrow B_t) | A_t \right) = \frac{(-b_t)^2}{A_t} = 0 \tag{155}$$

Proof.

In order to get clear on the Chi-square goodness-of fit test of a *sufficient condition*, we present here and point to a possible way out. In general, it is

$$+1 = +1 \tag{156}$$

In fact, it remains a matter of dispute whether this simple and to some extent cruel-looking axiom is so meaningful as it appears to be. But given that the relation in question is a matter of simple logic, this axiom can be the starting point of our further progress. Ultimately, for this reason, the equation before changes to

$$p(A_t \rightarrow B_t) \equiv p(a_t) + p(c_t) + p(d_t) = 1 \tag{157}$$

according to our definition of the *conditio per quam* relationship of a population. The point is powerful inasmuch as it lets us simplify the equation a

$$p(a_t) = 1 - (p(c_t) + p(d_t)) \tag{158}$$

or as

$$p(a_t) = 1 - (p(\underline{A}_t)) \tag{159}$$

As a matter of historical fact, the point is that the question of what makes this equation meaningful is the relationship $p(A_t) = 1 - p(\underline{A}_t)$. The exact nature of a *conditio per quam* relationship is easy to pinpoint and simplifies under the press of pure mathematics as

$$p(a_t) = p(A_t) \tag{160}$$

Facts as opposed to theories or values demand us that

$$N_t \times p(a_t) = N_t \times p(A_t) \tag{161}$$

Per definition it $a_t = N_t \times p(a_t)$ is or $A_t = N_t \times p(A_t)$. Judged solely in terms of mathematics, we obtain

$$a_t = A_t \tag{162}$$

or

$$a_t - A_t = 0 \tag{163}$$

The inward/outward looking nature of this equation doesn't change by the following operation.

$$(a_t - A_t)^2 = 0^2 \tag{164}$$

Dividing by A_t , it is

$$\frac{(a_t - A_t)^2}{A_t} = \frac{0}{A_t} = 0 \quad (165)$$

Most importantly of all, it is doubtful, however, that something which is a complex and self-organized whole might be reduced only to one single aspect. Indeed, many times, an attempt to reduce something complex to something simple can be associated with some shortcomings. There for, the whole aspect of a *conditio per quam* relationship should be considered. The equation before changes to

$$\frac{(a_t - A_t)^2}{A_t} + 0 = \frac{0}{A_t} + 0 = 0 \quad (166)$$

or to

$$\frac{(a_t - A_t)^2}{A_t} + \frac{((c_t + d_t) - (c_t + d_t))^2}{(c_t + d_t)} = \frac{0}{A_t} + 0 = 0 \quad (167)$$

or as

$$X^2 ((A_t \rightarrow B_t) | A_t) = \frac{(a_t - A_t)^2}{A_t} = 0 \quad (168)$$

More broadly, it is $A_t = a_t + b_t$. The exact nature of the equation before changes to

$$X^2 ((A_t \rightarrow B_t) | A_t) = \frac{(a_t - (a_t + b_t))^2}{A_t} = 0 \quad (169)$$

Finally, the equation before yields the Chi-square goodness of fit test of a *sufficient condition* without using the *continuity correction*, as

$$X^2 ((A_t \rightarrow B_t) | A_t) = \frac{(-b_t)^2}{A_t} = 0 \quad (170)$$

Quod erat demonstrandum.

Depending upon personal approach, study design or other factors, another equivalent method to calculate the chi-square value of a *conditio per quam* relationship with the *continuity correction* can be derived as demonstrated before as

$$X^2 ((A_t \rightarrow B_t) | B_t) \equiv \frac{((b_t) - (1/2))^2}{B_t} + 0 = 0 \quad (171)$$

A *fair study design* should assure that both methods provide the same chi-square value of a *conditio per quam* relationship.

Theorem. The Chi-square goodness-of-fit test of “Big Data”

There appears to be a growing interest in the analysis of “big datasets” with sizes beyond the ability of commonly used statistical methods. In this context, it is necessary to consider a variety of biases that are likely in the era of big data, including measurement error, sampling error, aggregation error, multiple comparisons errors, and other errors (Kaplan et al., 2014) associated with a systematic exclusion of information. At the heart of any scientific understanding of “big data” or large sample size studies which may integrate information from different data sources or/and from many thousands of persons is the notions of representation. “Big data” studies can be misleading and can lead to big inferential errors or are of a relatively little value if a large sample size of a study is not an accurate representative of the population to which the results of a study will be generalized. Even “big data” studies can go wrong and are not for sure more reliable than studies with smaller sample size (*The 1936 Literary Digest presidential election disaster* (Kaplan et al., 2014)) and it is more or less necessary to exercise greater caution to be able to rely on a big sample size. The *study design* should assure that a sample drawn from a population, a subset of a population, is an accurate and unbiased indication of what the population is like. A chi-square goodness-of-fit test proves whether an *observed* frequency distribution differs from a *theoretical* distribution.

Claim.

The chi-square goodness-of-fit test for an extremely very large sample follows as

$$X^2 = \frac{(\ln(\text{Observed}_t) - \ln(\text{Expected}_t))^2}{\ln(\text{Expected}_t)} \quad (172)$$

Proof.

For an *extremely large sample size*, the conventional chi-square goodness-of-fit test may perform very poorly even if the mathematical foundation of a very large sample is the same as for small data. In general, the mathematical foundation of a chi-square goodness-of-fit test is the equation

$$\text{Observed}_t = \text{Expected}_t \quad (173)$$

The large sample behavior of chi-square goodness-of-fit test improves especially for an extremely very large sample while taking the *natural logarithm*. The basic relationship doesn't change at all. It is

$$\ln(\text{Observed}_t) = \ln(\text{Expected}_t) \quad (174)$$

where *ln* denotes the *logarithmus naturalis*. Rearranging equation, it is

$$\ln(\text{Observed}_t) - \ln(\text{Expected}_t) = 0 \quad (175)$$

or

$$(\ln(\text{Observed}_t) - \ln(\text{Expected}_t))^2 = 0^2 \quad (176)$$

or

$$\frac{(\ln(\text{Observed}_t) - \ln(\text{Expected}_t))^2}{\ln(\text{Expected}_t)} = 0^2 \quad (177)$$

The chi-square goodness-of-fit test for an extremely very large sample follows

$$X^2 \equiv \frac{(\ln(\text{Observed}_t) - \ln(\text{Expected}_t))^2}{\ln(\text{Expected}_t)} \quad (178)$$

Quod erat demonstrandum.

For *preliminary reasons*, a sample size of $n = 1000$ and more can be treated as “Big Data”. The chi-square goodness-of-fit test for an extremely very large is an alternative to the *G-square test* (Quine and Robinson, 1985)

or to *Cramer's V test* et cetera. From a mathematical point of view, a chi-square goodness-of-fit test of a sufficient condition with a very large sample size n can be adopted to these new aspects too. Under these circumstances, it appears to be appropriate to take the logarithm of the data itself as

$$\ln(a_t) = \ln(A_t) \quad (179)$$

or as

$$\ln(a_t) - \ln(A_t) = 0 \quad (180)$$

and then implement the chi-square goodness-of-fit test for "Big Data" of a conditio per quam relationship on these transformed data as

$$X^2 \left((A_t \rightarrow B_t) | A_t \right) = \frac{(\ln(a_t) - \ln(A_t))^2}{\ln(A_t)} = 0 \quad (181)$$

Theorem. The Chi-square goodness-of-fit test of "Supra-Big Data"

Claims.

For „Supra-Big Data”, the Chi-square goodness-of-fit test is

$$X^2(Observed_t, Expected_t) = \frac{\left(\left(\ln \dots \left(\ln \left(\ln(Observed_t) \right) \right) \right) - \left(\ln \dots \left(\ln \left(\ln(Expected_t) \right) \right) \right) \right)^2}{\left(\ln \dots \left(\ln \left(\ln(Expected_t) \right) \right) \right)} = 0 \quad (182)$$

Proof.

Even for „Supra-Big Data” it stays valid that

$$Observed_t = Expected_t \quad (183)$$

Taking the *natural logarithm* we obtain

$$\ln(Observed_t) = \ln(Expected_t) \quad (184)$$

Repeating this operation one times it is

$$\ln \left(\ln(Observed_t) \right) = \ln \left(\ln(Expected_t) \right) \quad (185)$$

Even after repeating the operation before many times it is

$$\left(\ln \dots \left(\ln \left(\ln(Observed_t) \right) \right) \right) = \left(\ln \dots \left(\ln \left(\ln(Expected_t) \right) \right) \right) \quad (186)$$

where \ln denotes the *logarithmus naturalis*. The chi-square goodness-of-fit test a conditio per quam relationship of extremely "Supra-Big Data" samples follows as

$$X^2(Observed_t, Expected_t) = \frac{\left(\left(\ln \dots \left(\ln \left(\ln(Observed_t) \right) \right) \right) - \left(\ln \dots \left(\ln \left(\ln(Expected_t) \right) \right) \right) \right)^2}{\left(\ln \dots \left(\ln \left(\ln(Expected_t) \right) \right) \right)} = 0 \quad (187)$$

Quod erat demonstrandum.

Theorem. The Chi-square goodness-of fit test of an Exclusion relationship (antidote) I

Human medicine increasingly technology-dependent, can be characterized as a non-ending or everlasting administration of special antidotes to prevent an effect of something on something other. The truth is simply that many times human condition is not improving, either circumstances are getting worse as long as an useful antidote is not applied. We are forced to develop statistical methods that are jointly sufficient for a careful and systematic analysis of data which might provide us some help to *identify antidotes*. Thus far, let an event A_t be a sufficient condition of an event C_t or it is $\mathbf{p}(A_t \rightarrow C_t) = 1$. An event or an antidote B_t can counteract the effects of A_t with respect to C_t even if A_t is given. Let $p(A_t)$ denote the probability of a condition (i.e. a risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability of A_t and B_t . It is $1 \times 1 = 1$ and $\mathbf{p}(A_t \rightarrow C_t) \times \mathbf{p}(A_t | B_t) = 1$. Under conditions where the relationship between A_t and B_t is determined by an *exclusion relationship between A_t and B_t* it is $\mathbf{p}(A_t | B_t) = p(b_t) + (1 - p(A_t)) = p(c_t) + (1 - p(B_t)) = 1$.

Claim.

In general, a mathematical formula of the Chi-square goodness-of fit test of an *exclusion relationship* can be derived as

$$X^2 \left((A_t | B_t) | A_t \right) = \frac{(-a_t)^2}{A_t} = 0 \tag{188}$$

Proof.

What were some of the useful circumstances under which a Chi-square goodness-of fit test of an *exclusion relationship* can be derived. The truth is simply that

$$+1 = +1 \tag{189}$$

In fact, the equation before changes to

$$p(A_t | B_t) \equiv p(b_t) + p(c_t) + p(d_t) = 1 \tag{190}$$

according to our definition of the *exclusion relationship* of a population. Let us simplify the equation as

$$p(b_t) = 1 - (p(c_t) + p(d_t)) \tag{191}$$

or as

$$p(b_t) = 1 - (p(\underline{A}_t)) \tag{192}$$

As a matter of it is $\mathbf{p}(A_t) = 1 - \mathbf{p}(\underline{A}_t)$. One of the foundations of the *exclusion relationship* follows as

$$p(b_t) = p(A_t) \tag{193}$$

Multiplying by N_t , we obtain

$$N_t \times p(b_t) = N_t \times p(A_t) \tag{194}$$

Per definition it $b_t = N_t \times p(b_t)$ is or $A_t = N_t \times p(A_t)$. Thus far, in terms of mathematics, it is

$$b_t = A_t \tag{195}$$

or

$$b_t - A_t = 0 \tag{196}$$

Rearranging equation, it is

$$(b_t - A_t)^2 = 0^2 \tag{197}$$

Dividing by A_t , it is

$$\frac{(b_t - A_t)^2}{A_t} = \frac{0}{A_t} = 0 \quad (198)$$

The equation before changes to

$$\frac{(b_t - A_t)^2}{A_t} + 0 = \frac{0}{A_t} + 0 = 0 \quad (199)$$

or to

$$\frac{(b_t - A_t)^2}{A_t} + \frac{((c_t + d_t) - (c_t + d_t))^2}{(c_t + d_t)} = \frac{0}{A_t} + 0 = 0 \quad (200)$$

or as

$$X^2 \left((A_t \rightarrow | B_t) | A_t \right) = \frac{(b_t - A_t)^2}{A_t} + 0 = 0 \quad (201)$$

More broadly, it is $A_t = a_t + b_t$. The exact nature of the equation before changes to

$$X^2 \left((A_t | B_t) | A_t \right) = \frac{(b_t - (a_t + b_t))^2}{A_t} = 0 \quad (202)$$

Finally, the Chi-square goodness of fit test of an *exclusion relationship* without using *the continuity correction*, with degrees of freed d.f. = 2-1 = 1 follows as

$$X^2 \left((A_t | B_t) | A_t \right) = \frac{(-a_t)^2}{A_t} = 0 \quad (203)$$

Quod erat demonstrandum.

Theorem. The Chi-square goodness-of fit test of an Exclusion relationship (antidote) II

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability of A_t and B_t . Under conditions were the relationship between A_t and B_t is determined by an *exclusion relationship between A_t and B_t* it is $p(A_t | B_t) = p(b_t) + (1 - p(A_t)) = p(c_t) + (1 - p(B_t)) = 1$.

Claim.

In general, a mathematical formula of the Chi-square goodness-of fit test of an *exclusion relationship* can be derived as

$$X^2 \left((A_t | B_t) | B_t \right) = \frac{(-a_t)^2}{B_t} = 0 \quad (204)$$

Proof.

The Chi-square goodness-of fit test of an *exclusion relationship* can be derived from the equation

$$+1 = +1 \quad (205)$$

In fact, the equation before changes to

$$p(A_t | B_t) \equiv p(b_t) + p(c_t) + p(d_t) = 1 \quad (206)$$

according to our definition of the *exclusion* relationship of a population. Let us simplify the equation as

$$p(c_t) = 1 - (p(b_t) + p(d_t)) \quad (207)$$

or as

$$p(c_t) = 1 - (p(B_t)) \quad (208)$$

As a matter of it is $\mathbf{p(B_t)} = \mathbf{1 - p(B_t)}$. The *exclusion* relationship is determined as

$$p(c_t) = p(B_t) \quad (209)$$

Multiplying by N_t , we obtain

$$N_t \times p(c_t) = N_t \times p(B_t) \quad (210)$$

Per definition it $c_t = N_t \times p(c_t)$ or $B_t = N_t \times p(B_t)$. Thus far, in terms of mathematics, it is

$$c_t = B_t \quad (211)$$

or

$$c_t - B_t = 0 \quad (212)$$

Rearranging equation, it is

$$(c_t - B_t)^2 = 0^2 \quad (213)$$

Dividing by B_t , it is

$$\frac{(c_t - B_t)^2}{B_t} = \frac{0^2}{B_t} = 0 \quad (214)$$

The equation before changes to

$$\frac{(c_t - B_t)^2}{B_t} + 0 = \frac{0}{B_t} + 0 = 0 \quad (215)$$

or to

$$\frac{(c_t - B_t)^2}{B_t} + \frac{((b_t + d_t) - (b_t + d_t))^2}{(b_t + d_t)} = \frac{0}{B_t} + 0 = 0 \quad (216)$$

or as

$$X^2 \left((A_t | B_t) | B_t \right) = \frac{(c_t - B_t)^2}{B_t} + 0 = 0 \quad (217)$$

More broadly, it is $\mathbf{B_t} = \mathbf{a_t} + \mathbf{c_t}$. The exact nature of the Chi-square goodness-of fit test of *an exclusion relationship* changes to

$$X^2 \left((A_t | B_t) | B_t \right) = \frac{(c_t - (a_t + c_t))^2}{B_t} = 0 \quad (218)$$

Finally, the Chi-square goodness of fit test of an exclusion relationship without using the continuity correction, follows as

$$X^2 \left((A_t | B_t) | B_t \right) = \frac{(-a_t)^2}{B_t} = 0 \quad (219)$$

Quod erat demonstrandum.

Theorem. Self-contradictory data I

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability that A_t and B_t will occur/has occurred. Under conditions where the relationship between two random events abbreviated as A_t and B_t is determined by a necessary condition it is $p(A_t \leftarrow B_t) = p(A_t \text{ and } B_t) + (1 - p(B_t)) = 1$ and equally $\mathbf{p(A_t \text{ and } B_t) = p(B_t)}$.

Claim.

In general, under circumstances where $\mathbf{p(A_t) < 1}$ and $\mathbf{p(A_t \text{ and } B_t) = p(B_t)}$ it is

$$\begin{array}{lcl}
 k(A_t, B_t) & > & +0 \\
 & \mathbf{Proof.} & \\
 +1 & > & p(A_t) \\
 p(B_t) & > & p(B_t) \times p(A_t) \\
 \text{Since } p(A_t \leftarrow B_t) = p(a_t) + (1 - p(B_t)) = +1 \text{ it is then} & & p(a_t) = p(B_t). \text{ Substituting, we obtain} \\
 p(a_t) & > & (p(B_t) \times p(A_t)) \\
 & & (220) \\
 p(a_t) - (p(A_t) \times p(B_t)) & > & +0 \\
 \frac{p(a_t) - (p(A_t) \times p(B_t))}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} & > & \frac{0}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} \\
 k(A_t, B_t) & > & +0 \\
 & \mathbf{Q. e. d.} &
 \end{array}$$

Under conditions where $p(A_t) = 1$ it follows that $\mathbf{k(A_t \text{ and } B_t) = 0}$ and mathematically, A_t and B_t have to be treated as being independent of each other. In many problems, data gained from some observations provide an opportunity to increase the degree of confidence, when a decision is made to either accept the null hypothesis or accept the alternative hypothesis. Clearly, the null hypothesis and the alternative hypothesis are mutually exclusive thus that either the null hypothesis is false and the alternative hypothesis is true or the null hypothesis is true and the alternative hypothesis is false. In other words, a study design which provides data supporting the null-hypothesis: without A_t no B_t cannot at the same time support the hypothesis that $k < 0$. Such data are self-contradictory and cannot be used for further analysis.

Theorem. Self-contradictory data II

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability that A_t and B_t will occur/has occurred. Under conditions where the relationship between A_t and B_t is determined by an *exclusion relationship* $p(A_t | B_t)$ it is equally $p(A_t | B_t) = p(c_t) + (1 - p(B_t)) = p(b_t) + (1 - p(A_t)) = 1$ and $\mathbf{p(A_t \text{ and } B_t) = p(a_t) = 0}$.

Claim.

In general, an *exclusion* relationship demands that

$$k(A_t, B_t) < +0$$

Proof.

$$p(A_t | B_t) \equiv 1 - p(A_t) = 1$$

It is euqally

$$p(A_t \cap B_t) \equiv p(a_t) = 1 - 1 = 0$$

Furthermore, it is

$$k(A_t | B_t) = \frac{p(a_t) - (p(A_t) \times p(B_t))}{\sqrt{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}}$$

Under conditions where $p(A_t \cap B_t) \equiv p(a_t) = 0$ it is

(221)

$$k(A_t, B_t) = \frac{0 - (p(A_t) \times p(B_t))}{\sqrt{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}}$$

or

$$k(A_t, B_t) < 0$$

Q. e. d.

In other words, data which provide significant evidence that A_t excludes B_t and vice versa should equally demand that the causal relationship should be $k(A_t, B_t) < 0$, otherwise the data should be treated as self-contradictory.

Theorem. Self-contradictory data III

Let $p(A_t)$ denote the probability of the condition (i.e. risk factor), let $p(B_t)$ denote the probability of the conditioned (i.e. the outcome), let $p(A_t \text{ and } B_t)$ denote the joint probability of A_t and B_t . Under conditions where the relationship between A_t and B_t is determined by a *sufficient condition* it is $p(A_t \rightarrow B_t) = p(A_t \text{ and } B_t) + (1 - p(A_t)) = 1$ and it is equally $\mathbf{p(A_t \text{ and } B_t) = p(A_t)}$. In general, under circumstances where $\mathbf{p(B_t) < 1}$, it is

$$\begin{array}{rcl}
 k(A_t, B_t) & > & +0 \\
 & \textbf{Proof.} & \\
 +1 & > & p(B_t) \\
 p(A_t) & > & p(A_t) \times p(B_t) \\
 \text{Since } p(A_t \rightarrow B_t) = p(a_t) + (1 - p(A_t)) = +1 & \text{it is then} & p(a_t) = p(A_t). \text{Substituting, we obtain} \\
 p(a_t) & > & (p(A_t) \times p(B_t)) \tag{222} \\
 p(a_t) - (p(A_t) \times p(B_t)) & > & +0 \\
 \frac{p(a_t) - (p(A_t) \times p(B_t))}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} & > & \frac{0}{\sqrt[2]{p(A_t) \times (1 - p(A_t)) \times p(B_t) \times (1 - p(B_t))}} \\
 k(A_t, B_t) & > & +0 \\
 & \textbf{Q. e. d.} &
 \end{array}$$

Under conditions where $p(B_t) = 1$ it follows that $\mathbf{k(A_t \text{ and } B_t) = 0}$ and A_t and B_t must be treated as being independent of each other. In many problems, data gained from some observations provide an opportunity to increase the degree of confidence, when a decision is made to *either* accept the null hypotheses *or* accept the alternative hypothesis. Clearly, the null hypotheses and the alternative hypotheses are mutually exclusive thus that exactly one of the hypotheses must be true. Still, the quality of data varies and data as such do not assure an exact and true picture of reality with the consequence that a decision of an investigator can be wrong in principle.

An investigator can accept null hypotheses as true even if the same is wrong and vice versa. It is possible to accept alternative hypotheses as true even if the same is wrong. Data which provide evidence that A_t is a sufficient condition of B_t must not in the same respect provide evidence that there is a significant cause effect relationship. In fact, our ability to recognize conditions or risk factors might be seriously endangered by treating a cause as being identical with a condition. A cause is a condition too but not vice versa. A condition must not be a cause. Therefore, and due to mathematical requirements, *a significant cause effect relationship is not necessary to establish a significant sufficient condition relationship*. The analysis of alleged examples can show, among other things, how sufficient conditions should be understood, especially with relation to causation.

Example

For example there might be wet and dry conditions of a street while the relationship between raining and the state of a street is measured or investigated in a case control study. Rain (A_t) is generally known to be a sufficient condition for wet streets (B_t). In other words, rain as such guarantees that the event ‘the street is wet’ occurs. *If* it is raining *then* the street is wet ($n=1000$). Every time it is raining, the street gets wet, which was measured $n=4$ times. It isn’t raining and the street isn’t wet was documented $n=500$ times (*not raining and street not wet*). It is raining but the street wasn’t wet (*raining and street not wet*) was not measured all ($n=0$).

However, the presence of a street which is wet is not enough to conclude that was raining. In point of fact, there are also other possible factors (n=496) which are able to make the street wet (*not raining and street is wet*). The neighbor might have poured water on the street; a lorry may have lost oil et cetera. The data of this investigation are viewed by the table 5.

Table 5: The relationship between rain and a street

		The street is wet <B _t >		Total
		Yes	No	
It is raining <A _t >	Yes	4	b_t = 0	4
	No	496	500	996
Total		500	500	1000

$$k = +0.0634$$

$$p \text{ value } (k) = 0.06212481$$

The Chi-square goodness of fit test of a *sufficient condition* without using *the continuity correction*, is defined as

$$X^2 \left((A_t \rightarrow B_t) | \underline{B}_t \right) = \frac{(-b_t)^2}{N_t - B_t} = \frac{(-b_t)^2}{A_t} = X^2 \left((A_t \rightarrow B_t) | A_t \right) = 0 \quad (223)$$

According to the data above, we obtain

$$X^2 \left((A_t \rightarrow B_t) | \underline{B}_t \right) = \frac{(-0_t)^2}{500} = \frac{(-0_t)^2}{4} = X^2 \left((A_t \rightarrow B_t) | A_t \right) = 0 \quad (224)$$

Independent of study design, *both methods provide the same Chi square value*. The data agree with the null-hypothesis that the sample distribution does not differ from the theoretical population of the conditio per quam relationship. In other words, *if it is raining, then the street is wet*. Still, it is worth to mention that the data as obtained by the investigation presented before support the hypothesis too, that A_t and B_t are independent of each other (k = +0.0634, p value (k) = 0.06212481) which is a contradiction. Clearly, if it is raining then the street is wet with the consequence that the process of raining and a street which becomes wet cannot be treated as being independent of each other. Thus far, the conclusion drawn depend to a very great extent upon study design and other factors and *it is possible that data support the null-hypothesis: A_t is a sufficient condition of B_t while the causal relationship k is not significant. Such data are not self-contradictory and can be used for the analysis of conditions or risk factors, but not for causal analysis*. In particular, as proofed before, even under these circumstances it is necessary that the cause effect relationship should at least be greater than zero or **k > 0** otherwise the data are potentially self-contradictory and should not be used even for the analysis of conditions or risk factors.

Example.

A second study group investigated once again the relationship between the risk factor rain and the outcome “street is wet” and obtained the following data (Table 6).

Table 6: Antidot

		The street is wet 		
		Yes	No	Total
It is raining <A>	Yes	4	20	24
	No	496	480	976
Total		500	500	1000

$k = -0.1045$
 p value (k) = **0.00056237**
Odds ratio = 0.1935
 95% CI (Odds ratio) = (0.0657; 0.5704)
 IF <A> THEN
 p (IMP)= 0.9800
 X^2 (IMP)= **15.8438**

Even if the relationship between rain and the state of a street is clear, it is necessary to consider the following case. Different conditions of investigation can have an impact on the quality of conclusions based on data gained by studies. The data presented before do support a null hypothesis that *the rain has a protective effect against the wetness of a street* (Odds ratio = 0.1935; 95% CI (Odds ratio) = (0.0657; 0.5704)). It is *raining and the street isn't wet* was measured $n=20$ times wet (*raining and street not wet*) which contradicts our everyday experience. How is such a result possible? One reason for such a fundamental error can be an incorrect definition of cases and controls and an unfair study design. The index of unfairness is $IOU = (((24+500)/1000)-1) = -0.476$ and indicates *a very unfair study design*. Furthermore, it is possible that the street was wet but not recognized as being wet or not recorded as being wet although it rained. In other studies, the controls may have been contaminated et cetera. A mismatch of cases and controls excluded, it is possible that the control group possess an *antidote against the effect of the rain on the street*. In other words, it is possible that the measurements were performed under conditions where the street was protected against the effect of the rain i.e. by a great (transparent) tent or something similar thus that the street could not become wet even if it was possible to observe that it was raining. The conditions, inclusion and exclusion criteria et cetera under which investigations are performed can have fundamental influence on the quality of data and the validity of the conclusions drawn. Truth is one of the central subjects in scientific inquiry. And yet, despite a long history of debate in its own right going back for more than thousands of years the truth was, is and stays relative. Narrowly speaking, the truth or falsity of a scientific conclusion is based on many factors, among them the quality of data and the circumstances of investigation and has the potential to vary, sometimes extensively. In addition to a careful systematic observation and experiments, any scientific success achieved requires appropriate methods of scientific inference which enable us to infer beyond what is known by observation.

Theorem. Self-contradictory data IV

Claim.

The data are “suspicious” for significant bias or potentially self-contradictory if

$$X^2 \left((A_t \rightarrow B_t) | (N_t - B_t) \right) < X^2 \left((A_t \rightarrow B_t) | (A_t) \right) \quad (225)$$

Proof.

As long as the whole population is not investigated, the study design of a case-control or of another study should assure that the same chi square value can be achieved from the data recorded. This condition is seldom provided by studies published. Many times, the study design demands or assures conditions or a sample were

$$A_t < (N_t - B_t) \quad (226)$$

Multiplying by $(-b)^2/(A_t \times (N_t - B_t))$ it is

$$A_t \times \left(\frac{(-b_t)^2}{A_t \times (N_t - B_t)} \right) < (N_t - B_t) \times \left(\frac{(-b_t)^2}{A_t \times (N_t - B_t)} \right) \quad (227)$$

Simplifying, we obtain

$$\left(\frac{(-b_t)^2}{(N_t - B_t)} \right) < \left(\frac{(-b_t)^2}{A_t} \right) \quad (228)$$

or

$$X^2 \left((A_t \rightarrow B_t) | (N_t - B_t) \right) < X^2 \left((A_t \rightarrow B_t) | (A_t) \right) \quad (229)$$

Quod erat demonstrandum.

Thus far, a study design which demands or assures that $A_t < (N_t - B_t)$ can lead to biased Chi-square values characterized by the formula $X^2(A_t \rightarrow B_t | (N_t - B_t)) < X^2(A_t \rightarrow B_t | (A_t))$ and the question arises, which X^2 should be used? Statistical tests primary handle samples and not populations. Still, an appropriate sample should assure that something insignificant stays significant and that statistical tests correctly applied should have the same chance to rejects a false null hypothesis.

Theorem. Self-contradictory data V

Claim.

The data are “suspicious” for significant bias or potentially self-contradictory if

$$X^2 \left((A_t \rightarrow B_t) | (N_t - B_t) \right) > X^2 \left((A_t \rightarrow B_t) | (A_t) \right) \quad (230)$$

Proof.

Circumstances were the study design demands or assures conditions or a sample were

$$A_t > (N_t - B_t) \quad (231)$$

are leading to similar point of view. Multiplying by $(-b)^2/(A_t \times (N_t - B_t))$ it is

$$A_t \times \left(\frac{(-b_t)^2}{A_t \times (N_t - B_t)} \right) > (N_t - B_t) \times \left(\frac{(-b_t)^2}{A_t \times (N_t - B_t)} \right) \quad (232)$$

Simplifying, we obtain

$$\left(\frac{(-b_t)^2}{(N_t - B_t)} \right) > \left(\frac{(-b_t)^2}{A_t} \right) \quad (233)$$

or

$$X^2 \left((A_t \rightarrow B_t) | (N_t - B_t) \right) > X^2 \left((A_t \rightarrow B_t) | (A_t) \right) \quad (234)$$

Quod erat demonstrandum.

Again, a study design which is grounded on the assumption that $A_t > (N_t - B_t)$ leads to the conclusion that $X^2(A_t \rightarrow B_t | (N_t - B_t)) > X^2(A_t \rightarrow B_t | (A_t))$ and the question arises again, which X^2 is valid and which X^2 should be used.

Theorem. Self-contradictory data VI

Claim.

The data are “suspicious” for significant bias or potentially self-contradictory if

$$X^2 \left((A_t \leftarrow B_t) | (N_t - B_t) \right) < X^2 \left((A_t \leftarrow B_t) | (A_t) \right) \quad (235)$$

Proof.

As long as the whole population is not investigated, the study design of a case-control or of another study should assure that the same chi square value can be achieved from the data recorded. This condition is seldom provided by studies published. Many times, the study design demands or assures conditions or as sample were

$$B_t < (N_t - A_t) \quad (236)$$

Multiplying by $(-c)^2 / (B_t \times (N_t - A_t))$ it is

$$B_t \times \left(\frac{(-c_t)^2}{B_t \times (N_t - A_t)} \right) < (N_t - A_t) \times \left(\frac{(-c_t)^2}{B_t \times (N_t - A_t)} \right) \quad (237)$$

Simplifying, we obtain

$$\left(\frac{(-c_t)^2}{(N_t - A_t)} \right) < \left(\frac{(-c_t)^2}{B_t} \right) \quad (238)$$

or

$$X^2 \left((A_t \leftarrow B_t) | (N_t - A_t) \right) < X^2 \left((A_t \leftarrow B_t) | (B_t) \right) \quad (239)$$

Quod erat demonstrandum.

Thus far, a study design which demands that $A_t < (N_t - B_t)$ leads to a Chi-square values which are characterized by the formula $X^2(A_t \leftarrow B_t | (N_t - A_t)) < X^2(A_t \leftarrow B_t | (B_t))$ and the question arises, which X^2 is the correct one.

Theorem. Self-contradictory data VII

Claim.

The data are “suspicious” for significant bias or potentially self-contradictory (from the standpoint of a necessary condition) if

$$X^2 \left((A_t \leftarrow B_t) | (N_t - B_t) \right) > X^2 \left((A_t \leftarrow B_t) | (A_t) \right) \quad (240)$$

Proof.

A study design which demands or assures conditions or as sample were

$$B_t > (N_t - A_t) \quad (241)$$

can be misleading too. Multiplying by $(-c)^2 / (B_t \times (N_t - A_t))$ it is

$$B_t \times \left(\frac{(-c_t)^2}{B_t \times (N_t - A_t)} \right) > (N_t - A_t) \times \left(\frac{(-c_t)^2}{B_t \times (N_t - A_t)} \right) \quad (242)$$

Simplifying, we obtain

$$\left(\frac{(-c_t)^2}{(N_t - A_t)} \right) > \left(\frac{(-c_t)^2}{B_t} \right) \quad (243)$$

or

$$X^2 \left((A_t \leftarrow B_t) | (N_t - A_t) \right) > X^2 \left((A_t \leftarrow B_t) | (B_t) \right) \quad (244)$$

Quod erat demonstrandum.

Again, a study design which is grounded on the assumption that $A_t > (N_t - B_t)$ leads to $X^2(A_t \leftarrow B_t | (N_t - B_t)) > X^2(A_t \leftarrow B_t | (A_t))$ and the question arises again, which X^2 is valid and should be used.

Theorem. A fair study design I

Claim.

A study design from the standpoint of a *conditio per quam relationship* is fair and the data are formally not self-contradictory due to study design if

$$N_t = A_t + B_t \quad (245)$$

Proof.

The Chi-square value of a *conditio per quam relationship* is demands that

$$X^2 \left((A_t \rightarrow B_t) | \underline{B_t} \right) = \frac{(-b_t)^2}{N_t - B_t} = \frac{(-b_t)^2}{A_t} = X^2 \left((A_t \rightarrow B_t) | \underline{A_t} \right) \quad (246)$$

Both methods applied on a data body should yield the same Chi-square value. In other words, it is

$$\frac{(-b_t)^2}{N_t - B_t} = \frac{(-b_t)^2}{A_t} \quad (247)$$

Define $(-b_t)^2 \equiv 1$ and rearrange equation, it is

$$N_t - B_t = A_t \quad (248)$$

or

$$N_t = A_t + B_t \quad (249)$$

Quod erat demonstrandum.

Theorem. A fair study design II

Claim.

A study design from the standpoint of a *conditio sine qua non relationship* is fair and the data are formally not self-contradictory due to study design if

$$N_t = A_t + B_t \quad (250)$$

Proof.

The Chi-square value of a *conditio sine qua non relationship* demands that

$$X^2 \left((A_t \leftarrow B_t) | A_t \right) = \frac{(-c_t)^2}{N_t - A_t} = \frac{(-c_t)^2}{B_t} = X^2 \left((A_t \leftarrow B_t) | B_t \right) \quad (251)$$

Both methods applied on a data body should yield the same Chi-square value. In other words, it is

$$\frac{(-c_t)^2}{N_t - A_t} = \frac{(-c_t)^2}{B_t} \quad (252)$$

Define $(-c)^2 \equiv 1$ and rearrange equation, it is

$$N_t - A_t = B_t \quad (253)$$

or

$$N_t = A_t + B_t \quad (254)$$

Quod erat demonstrandum.

Theorem. A fair study design III

The guarantee of a fair study design is fundamental in any empirical scientific research and of every modern medical investigation. The framework of a fair study design should obey especially *the principle of equality of arms* which is a central feature of every scientific combat to ensure completely only the discovery of the truth. The *principle of equality of arms* leaves no room for defending material interest, ideological position or wishful thinking but requires that *advocates of a special null hypothesis* and *opponents of the same null hypotheses* have the same chance or possibilities to reject or to accept the null-hypothesis at their disposal. One could sum up the principle of equality of (scientific) arms by saying that *no party should have an unfair advantage over the other party especially due to study design*. Put in other terms, any scientific research is not complete without the notion of fairness. Ignoring the historical origins and theoretical foundations of the principle of equality of (scientific) arms a fair and careful study design directed to the goal that *a correct null-hypothesis has to be accepted* and that *a false null-hypothesis has to be rejected* is the core of evaluations to determine how believable a hypothesis is. Independently of the extent of the data to be recorded or the type of the study (case-control study, cohort study et cetera), formally, the design of the study must ensure that the analyzing results of the data generated are the same. In the following, this problem will be analyzed from the standpoint of the research on secondary data (i. e. case control studies) and the research on primary data **under ideal conditions** (No bias, no systematic errors, perfect accuracy of a measuring instrument et cetera). Aside from the type of study, we intended to give the same answer to scientific questions and to gain the same new knowledge.

Claim.

A study design which demands that $a_t = d_t$ is fair and the data are formally not self-contradictory due to study design if

$$A_t + B_t = N_t \tag{255}$$

Proof.

Sometimes, study design demands or assures conditions or a sample were

$$a_t = d_t \tag{256}$$

where a_t denotes the number of subjects (**exposed and diseased**) while d_t denotes the number of subjects (**not exposed and not diseased**). Adding $(b_t + c_t + d_t)$ to the equation before, it is

$$(a_t) + ((b_t) + (c_t) + (d_t)) = (d_t) + ((b_t) + (c_t) + (d_t)) \tag{257}$$

or

$$(a_t) + (b_t) + (c_t) + (d_t) = ((c_t) + (d_t)) + ((b_t) + (d_t)) \tag{258}$$

In general, it is $N_t = (a_t + b_t + c_t + d_t)$. Furthermore, it is $\underline{A}_t = (c_t + d_t)$ and $\underline{B}_t = (b_t + d_t)$. The equation changes to

$$N_t = \underline{A}_t + \underline{B}_t \tag{259}$$

In other words, under these conditions, the study design demands equally that

$$N_t - \underline{A}_t = \underline{B}_t \tag{260}$$

It is $A_t = (N_t - \underline{A}_t)$ and the equation before simplifies as

$$A_t = \underline{B}_t \tag{261}$$

It is $\underline{A}_t = (N_t - A_t)$ and $\underline{B}_t = (N_t - B_t)$. The equation above derived as

$$N_t = \underline{A}_t + \underline{B}_t \tag{262}$$

simplifies as

$$N_t = (N_t - A_t) + (N_t - B_t) \tag{263}$$

or to

$$N_t - (N_t - A_t) = (N_t - B_t) \quad (264)$$

or to

$$A_t = N_t - B_t \quad (265)$$

The condition $a_t = d_t$ is assured if

$$A_t + B_t = N_t \quad (266)$$

Quod erat demonstrandum.

Theorem. A fair study design IV

Claim.

A study design from the standpoint of an *exclusion relationship* is fair and the data are formally not self-contradictory due to study design if

$$N_t = \underline{A}_t + B_t \quad (267)$$

Proof.

The Chi-square value of an *exclusion relationship* demands that

$$X^2 \left((A_t | B_t) | A_t \right) = \frac{(-a_t)^2}{A_t} = \frac{(-a_t)^2}{B_t} = X^2 \left((A_t | B_t) | B_t \right) \quad (268)$$

Both methods applied on a data body should yield the same Chi-square value. In other words, it is

$$\frac{(-a_t)^2}{A_t} = \frac{(-a_t)^2}{B_t} \quad (269)$$

Define $(-a_t)^2 \equiv 1$ and rearrange equation, it is

$$A_t = B_t \quad (270)$$

or

$$N_t - A_t = N_t - B_t \quad (271)$$

or

$$\underline{A}_t = \underline{B}_t \quad (272)$$

The study design to test the exclusion relationship demands that $A_t = B_t$ and that $\underline{A}_t = \underline{B}_t$. This circumstances can be identical with the demand that $A_t = B_t = \underline{A}_t = \underline{B}_t$ but must not. In other words, an *exclusion relationship* demands a study design were $A_t = B_t$. Adding B_t it is

$$A_t + \underline{A}_t = B_t + \underline{A}_t \quad (273)$$

or

$$N_t = \underline{A}_t + B_t \quad (274)$$

Quod erat demonstrandum.

Theorem. A fair study design V

Claim.

A study design from the standpoint of *conditio sine qua non* and *conditio per quam* and an *exclusion relationship* is fair and the data are formally not self-contradictory due to study design if

$$N_t = \underline{A}_t + B_t \quad (275)$$

Proof.

The Chi-square value of an *exclusion relationship* demands that

$$N_t = \underline{A}_t + B_t \quad (276)$$

A study design from the standpoint of *conditio sine qua non* and *conditio per quam* relationship can be regarded as fair if $N_t = A_t + B_t$. Substituting this relationship into equation before, it is

$$A_t + B_t = \underline{A}_t + B_t \quad (277)$$

or

$$A_t = \underline{A}_t \quad (278)$$

At the same time, the study design should be fair with respect to an *exclusion relationship*. In this case, it is equally true that $A_t = B_t$. We obtain

$$A_t = \underline{A}_t = B_t = \underline{B}_t \quad (279)$$

Quod erat demonstrandum.

Theorem. A fair study design VI

Claim.

A study design which investigates the causal relationship between A_t and B_t should respect especially the law of independence. Whether the absence of independence may be or not be one aspect of the causal relationship or not, study design should ensure that an independence of A_t from B_t and vice versa can be recognized. Under the assumption of independence of A_t from B_t study design is fair and the data are formally not self-contradictory due to study design if

$$A_t = B_t \quad (280)$$

Proof.

Under the assumption of independence, it is

$$p(A_t \cap B_t) \equiv p(a_t) = p(A_t) \times p(B_t) \quad (281)$$

Multiplying by $N_t \times N_t$, it is

$$N_t \times N_t \times p(a_t) = N_t \times N_t \times p(A_t) \times p(B_t) \quad (282)$$

or

$$\begin{array}{ccc} N_t & \times & a_t \\ & / & \backslash \\ & & A_t \times B_t \end{array} \quad (283)$$

Rearranging equation, we obtain

$$\frac{N_t \times a_t}{A_t} = B_t \quad \text{and} \quad \frac{N_t \times a_t}{B_t} = A_t \quad (284)$$

or

$$\left(\left(\frac{N_t \times a_t}{A_t} \right) - B_t \right) = 0 \quad \text{and} \quad \left(\left(\frac{N_t \times a_t}{B_t} \right) - A_t \right) = 0 \quad (285)$$

or

$$\left(\left(\frac{N_t \times a_t}{A_t} \right) - \frac{A_t \times B_t}{A_t} \right) = 0 \quad \text{and} \quad \left(\left(\frac{N_t \times a_t}{B_t} \right) - \frac{A_t \times B_t}{B_t} \right) = 0 \quad (286)$$

or that

$$\left(\frac{(N_t \times a_t) - (A_t \times B_t)}{A_t} \right) = 0 \quad \text{and} \quad \left(\frac{(N_t \times a_t) - (A_t \times B_t)}{B_t} \right) = 0 \quad (287)$$

In other words, under the condition of independence, study design should fulfill the requirement that

$$\left(\frac{(N_t \times a_t) - (A_t \times B_t)}{A_t} \right) = \left(\frac{(N_t \times a_t) - (A_t \times B_t)}{B_t} \right) \quad (288)$$

We define $((N_t \times a_t) - (A_t \times B_t)) = 1$. The equation before simplifies to

$$\left(\frac{1}{A_t}\right) = \left(\frac{1}{B_t}\right) \quad (289)$$

Under the assumption of independence, a study design is fair if

$$A_t = B_t \quad (290)$$

Quod erat demonstrandum.

Example.

Under the assumption of a conditio sine qua non relationship or of a conditio per quam relationship study design should assured that

$$N_t = A_t + B_t \quad (291)$$

Under the condition of independence, it is at the same time $A_t = B_t$. Substituting we obtain

$$N_t = A_t + B_t = 2 \times A_t = 2 \times B_t \quad (292)$$

or

$$A_t = B_t = \left(\frac{N_t}{2}\right) \quad (293)$$

Theorem. Index of unfairness.

Aside from personnel, financial, organizational and logistical questions of a study, the scientific value of a (medical) study is determined especially by factors like study design, statistical methodology used, sample size calculations and a properly selected, highly representative study population with defined and selective inclusion and exclusion criteria. The extent to which the measuring technique and instruments used consistently provide the same results if measurements are repeated should be accurate enough. The significance of study design for the quality of the conclusions drawn is often underestimated. In point of fact, errors in the statistical evaluation can be corrected after the study has been completed. In contrast to errors in the statistical evaluation it is difficult to correct errors in study design afterwards. A number of potential problems may bias the results of observational studies or even or well-planned, experimental randomized clinical trials. Nevertheless, even if many questions in human medicine can only be answered with observational studies medical research studies itself provide already additional statistical information whether the data of a study can be considered to be evaluated by statistical test procedures to generalize the results from the sample for the whole population. The relation between data and hypotheses is of key importance in almost all empirical research and those who plan to perform a study which should make an important contribution to medical knowledge must occupy themselves intensively with an appropriate and careful study design. Statistical methods which are relating hypotheses in the light of empirical facts may enable us even to extrapolate from data to predictions and general facts. Some of the main methodological problems can be avoided if the foundations of statistical methods are logically and mathematically correct. Data have an impact on a hypothesis, but the impact should depend on the data themselves and not just on the study design of the researcher. The underlying question arises, therefore, how can such a problem (even *ex post*) be operationalized, meaning that it must be converted into an evaluable and measurable form.

Claim.

The index of unfairness (IOU) can be derived as

$$IOU = \left(\left(\frac{A_t + B_t}{N_t} \right) \right)^{-1} \tag{294}$$

Proof.

Under conditions were the data of a study are analyzed by a chi-square goodness of fit test of a necessary condition, the study design should assure, that the same chi square value should be achieved. In other words, it is

$$X^2 \left((A_t \leftarrow B_t) | \underline{A}_t \right) \equiv X^2 \left((A_t \leftarrow B_t) | B_t \right) \tag{295}$$

or

$$\frac{(-c_t)^2}{\underline{A}_t} \equiv \frac{(-c_t)^2}{B_t} \tag{296}$$

If $c_t = 0$, we set $c_t = 1$. We obtain

$$\frac{1}{\underline{A}_t} \equiv \frac{1}{B_t} \tag{297}$$

or

$$\underline{A}_t \equiv B_t \tag{298}$$

It is $\underline{A}_t = N_t - A_t$. Substituting this relationship into the equation before, we obtain

$$N_t - A_t \equiv B_t \tag{299}$$

or

$$N_t \equiv (A_t + B_t) \tag{300}$$

or

$$\frac{N_t}{N_t} = +1 \equiv \frac{(A_t + B_t)}{N_t} \quad (301)$$

The index of unfairness, abbreviated as IOU, follows as

$$IOU \equiv \left(\left(\frac{(A_t + B_t)}{N_t} \right) - 1 \right) = 0 \quad (302)$$

Q. e. d.

The range of *the index of unfairness* is [-1;+1]. In this context let IOF=0 denote a *fair* study design, let $0 < |\text{IOF}| \leq 0.25$ denote an *unfair* study design, let $0.25 < |\text{IOF}| \leq 0.5$ denote a *very unfair* study design, let $0.5 < |\text{IOF}| \leq 0.75$ denote a *highly unfair* study design and let $0.75 < |\text{IOF}| \leq 1$ denote an *extremely unfair* study design.

4. Discussion

Failure to apply rigorous standards and appropriate statistical methods in biomedical research on data collected from a valid scientific design can lead to misleading conclusions and the costs to patients and society could be high. Briefly, many times a very large sample is required to discover a small difference. Still, the sample size of medical studies is often too small, thus that the power is also too small and a relationship is either only described imprecisely (Moher et al., 1994) or even unidentified. Anyone who denies that a very careful calculation of the sample size of a study is one of the problems of scientific research may characteristically insist that the measurements itself may be invalid or false and may lead to erroneous conclusions too. Therefore, whatever the sample size calculation and other potential sources of bias, the selection of an appropriate study design, free of publication bias, is important too, to be able to generalize the study results to the population. *Funnel plots* are one of the several attempts made to assess the magnitude of publication bias. Many meta-analyses show Funnel plots as proposed in 1984 by Light and Pillemer (Light and Pillemer, 1984) to examine whether there is evidence against or for the presence of publication bias. The funnel plot is a kind of a scatter plot with some measure of weight (such as the sample size, the inverse variance, the standard error et cetera) on the vertical axis and the treatment effect on the horizontal axis. More precisely, the Funnel plot's (plots of effect estimates against sample size) wide popularity followed an article published in the BMJ in 1997 (Egger et al., 1997) is not unrestrictedly justified. The capacity of Funnel plot to detect publication bias in meta-analyses is often misleading (Lau et al, 2006) and equally inaccurate especially for meta-analyses of proportion studies (Hunter et al., 2014) with low proportion outcomes. To date, funnel plot may overlook serious bias (Lau et al, 2006) and it is still clear whether funnel plots really diagnose publication bias (Zwetsloot et al., 2017) at all. Formal statistical tests are sometimes useful tools to eliminate the subjectivity in scientific research. Nonetheless, bias starts when subjectivism begins. Any use and abuse of methodological tests depends on the chance of a study of being published too. Medical studies with results different from the published ones are often more endangered and threatened of not being published or have a decreased likelihood of being published. Medical studies with favorable results are published more often, and more quickly, than trials with negative findings which can lead to publication bias (Hopewell et al., 2009). As a result of such a scientific practice or scientific evolution or a natural (peer-reviewed dominated) selection, main-stream compatible or source of funding adequate favorable results are published more often than other ones. On the long run, there is a serious overestimation of the effects in the literature found and a damaging and escalating effect on the integrity of scientific knowledge is possible. Unfortunately, there is evidence suggesting that this systematic publication bias documented in the literature for decades is increasing (Joober et al., 2012). Finally, a prevention of publication bias is of course much more desirable than a corrective or diagnostic analysis or finding or excluding of publication bias. With regard to the prevention of publication bias *the index of unfairness* is of use since the same reduces errors due to study design. Nonetheless, the knowledge of publication bias even *ex-post* due to index of unfairness possible needed to detect or to avoid inconsistent conclusions too. The index of unfairness and other methods developed in this publication are useful tools for further evaluation of publication bias and can help to reduce the impact of publication bias on the certainty of scientific evidence.

5. Conclusion

The several attempts to assess the magnitude of publication bias require several assumptions which are difficult to ascertain and are labeled with various other limitations. The presence of publication bias can be determined even *ex post* while using the index of unfairness.

Conflict of interest statement

The authors declare that no conflict of interest exists according to the guidelines of the International Committee of Medical Journal Editors.

Acknowledgements

The public domain software program GnuPlot was used to draw the figures in this publication.

References

- Altman, D. G. (1999). *Practical statistics for medical research*. Boca Raton, Fla: Chapman & Hall/CRC.
- Arbuthnott, J. (1710). An Argument for Divine Providence, Taken from the Constant Regularity Observ'd in the Births of Both Sexes. By Dr. John Arbuthnott, Physitian in Ordinary to Her Majesty, and Fellow of the College of Physitians and the Royal Society. *Philosophical Transactions of the Royal Society of London*, 27(325–336), 186–190. <https://doi.org/10.1098/rstl.1710.0011>
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, 533(7604), 452–454. <https://doi.org/10.1038/533452a>
- Barukčić, I. (1989). *Die Kausalität* (1. Aufl.). Hamburg: Wiss.-Verl.
- Barukčić, I. (1997). *Die Kausalität* (2., völlig überarb. Aufl.). Wilhelmshaven: Scientia.
- Barukčić, I. (2005). *Causality: New statistical methods*. Norderstedt, Germany: Books on Demand GmbH.
- Barukčić, I. (2006a). *Causality: New statistical methods* (2. Aufl.). Norderstedt: Books on Demand.
- Barukčić, I. (2006b). *New Method for Calculating Causal Relationships*. Retrieved from <http://www.medicine.mcgill.ca/epidemiology/hanley/ibc2006program.pdf>
- Barukčić, I. (2009a). *Causality I. A theory of energy, time and space* (5. ed., 14. rev).
- Barukčić, I. (2009b). *Causality II. A theory of energy, time and space* (5. ed., 13. rev).
- Barukčić, I. (2011a). Anti Heisenberg-Refutation Of Heisenberg's Uncertainty Relation: In: ADVANCES IN QUANTUM THEORY: Proceedings of the International Conference on Advances in Quantum Theory, Växjö, (Sweden), 14–17 June 2010. In *American Institute of Physics - Conference Proceedings* (Vol. 1327, pp. 322–325). <https://doi.org/10.1063/1.3567453>
- Barukčić, I. (2011b). The Equivalence of Time and Gravitational Field. *Physics Procedia*, 22, 56–62. <https://doi.org/10.1016/j.phpro.2011.11.008>
- Barukčić, I. (2012). Anti-Bell - Refutation of Bell's theorem: In: Quantum Theory: Reconsideration of Foundations-6 (QTRF6), Växjö, (Sweden), 11-14 June 2012. In *American Institute of Physics - Conference Proceedings* (Vol. 1508, pp. 354–358). <https://doi.org/10.1063/1.4773147>
- Barukčić, I. (2016a). The Mathematical Formula of the Causal Relationship k. *International Journal of Applied Physics and Mathematics*, 6(2), 45–65. <https://doi.org/10.17706/ijapm.2016.6.2.45-65>
- Barukčić, I. (2016b). The Physical Meaning of the Wave Function. *Journal of Applied Mathematics and Physics*, 04(06), 988–1023. <https://doi.org/10.4236/jamp.2016.46106>
- Barukčić, I. (2016c). Unified Field Theory. *Journal of Applied Mathematics and Physics*, 04(08), 1379–1438. <https://doi.org/10.4236/jamp.2016.48147>
- Barukčić, I. (2017a). Helicobacter pylori—The Cause of Human Gastric Cancer. *Journal of Biosciences and Medicines*, 05(02), 1–9. <https://doi.org/10.4236/jbm.2017.52001>
- Barukčić, I. (2017b). *Theoriae causalitatis principia mathematica*. Norderstedt: Books on Demand.
- Barukčić, I. (2018a). Epstein Bar Virus—The Cause of Hodgkin's Lymphoma. *Journal of Biosciences and Medicines*, 06(01), 75–100. <https://doi.org/10.4236/jbm.2018.61008>
- Barukčić, I. (2018b). Fusobacterium nucleatum —The Cause of Human Colorectal Cancer. *Journal of Biosciences and Medicines*, 06(03), 31–69. <https://doi.org/10.4236/jbm.2018.63004>
- Barukčić, I. (2018c). Gastric Cancer and Epstein-Barr Virus Infection. *Modern Health Science*, 1(2), 1–18. <https://doi.org/10.30560/mhs.v1n2p1>
- Barukčić, I. (2018d). Helicobacter Pylori is the Cause of Gastric Cancer. *Modern Health Science*, 1(1), 43–50. <https://doi.org/10.30560/mhs.v1n1p43>
- Barukčić, I. (2018e). Human Cytomegalovirus is the Cause of Glioblastoma Multiforme. *Modern Health Science*, 1(2), 19. <https://doi.org/10.30560/mhs.v1n2p19>
- Barukčić, I. (2018f). Human Papillomavirus—The Cause of Human Cervical Cancer. *Journal of Biosciences and Medicines*, 06(04), 106–125. <https://doi.org/10.4236/jbm.2018.64009>

- Barukčić, I. (2018g). Mycobacterium Avium Subspecies Paratuberculosis: The Cause Of Crohn's Disease. *Modern Health Science*, 1(1), 19–34. <https://doi.org/10.30560/mhs.v1n1p19>
- Barukčić, I. (2018h). Mycobacterium Avium Subspecies Paratuberculosis: The Cause Of Crohn's Disease. *Modern Health Science*, 1(1), 19–34. <https://doi.org/10.30560/mhs.v1n1p19>
- Barukčić, I. (2018i). Epstein-Barr virus is the cause of multiple sclerosis. *International Journal of Current Medical and Pharmaceutical Research*, 4(9 (A)), 3674–3682. <http://dx.doi.org/10.24327/23956429.ijcmpr20180538>
- Barukčić, K., Barukčić, J. P., & Barukčić, I. (2018). Epstein-Barr virus is the cause of rheumatoid arthritis. *Romanian Journal of Rheumatology*, 27(4), 148–163.
- Bernoulli, J. (1713). *Ars conjectandi, Opus posthumus: Accedit Tractatus de seriebus infinitis; et epistola Gallice scripta De Ludo Pilae Reticularis*. Basileae (Basel, Suisse): Impensis Thurnisiorum [Tournes], fratrum. <https://doi.org/10.3931/e-rara-9001>
- Bienaymé, I.-J. (1846). Sur les probabilités des erreurs d'après la méthode des moindres carrés. *Journal De Mathématiques Pures Et Appliquées*, 1(17), 33–78. <https://doi.org/10.1001/jama.1983.03330370053031>
- Bortkiewicz, L. (Ed.). (1898). *Das Gesetz der kleinen Zahlen: [Transl. into English: The law of small numbers]*. Leipzig (Germany): B.G. Teubner. Retrieved from <https://ia802308.us.archive.org/33/items/dasgesetzderklei00bortrich/dasgesetzderklei00bortrich.pdf>
- Conover, W. J. (1974). Some Reasons for Not Using the Yates Continuity Correction on 2×2 Contingency Tables. *Journal of the American Statistical Association*, 69(346), 374–376. <https://doi.org/10.1080/01621459.1974.10482957>
- Cornfield, J. (1951). A method of estimating comparative rates from clinical data; applications to cancer of the lung, breast, and cervix. *Journal of the National Cancer Institute*, 11(6), 1269–1275.
- DeGroot, M. H., Schervish, M. J., Fang, X., Lu, L., & Li, D. (2005). *Probability and Statistics* (Third Edition [Rep. & arr. ed.]). Beijing (China): Higher Education Press.
- Dorey, F. (2010). In Brief: The P Value: What Is It and What Does It Tell You? *Clinical Orthopaedics and Related Research*, 468(8), 2297–2298. <https://doi.org/10.1007/s11999-010-1402-9>
- Edwards, A. W. F. (1963). The Measure of Association in a 2 × 2 Table. *Journal of the Royal Statistical Society. Series A (General)*, 126(1), 109. <https://doi.org/10.2307/2982448>
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, 315(7109), 629–634. <https://doi.org/10.1136/bmj.315.7109.629>
- Fisher, R. A. (1922). On the Interpretation of χ^2 from Contingency Tables, and the Calculation of P. *Journal of the Royal Statistical Society*, 85(1), 87. <https://doi.org/10.2307/2340521>
- Fisher, R. A. (1935). The Logic of Inductive Inference. *Journal of the Royal Statistical Society*, 98(1), 39. <https://doi.org/10.2307/2342435>
- Fisher, Ronald A. (1925). *Statistical Methods for Research Workers*. Edinburgh: Oliver and Boyd. Retrieved from http://www.haghigh.com/resources/materials/Statistical_Methods_for_Research_Workers.pdf
- Gomes, G. (2009). Are Necessary and Sufficient Conditions Converse Relations. *Australasian Journal of Philosophy*, 87, 375–387. <https://doi.org/10.1080/00048400802587325>
- Gonin, H. T. (1936). XIV. The use of factorial moments in the treatment of the hypergeometric distribution and in tests for regression. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 21(139), 215–226. <https://doi.org/10.1080/14786443608561573>
- Grimes, D. A., & Schulz, K. F. (2002). Bias and causal associations in observational research. *Lancet (London, England)*, 359(9302), 248–252. [https://doi.org/10.1016/S0140-6736\(02\)07451-2](https://doi.org/10.1016/S0140-6736(02)07451-2)
- Grizzle, J. E. (1967). Continuity Correction in the χ^2 -Test for 2 × 2 Tables. *The American Statistician*, 21(4), 28. <https://doi.org/10.2307/2682103>
- Hanley, J. A. (1983). If Nothing Goes Wrong, Is Everything All Right? *JAMA*, 249(13), 1743.
- Hopewell, S., Loudon, K., Clarke, M. J., Oxman, A. D., & Dickersin, K. (2009a). Publication bias in clinical trials due to statistical significance or direction of trial results. *Cochrane Database of Systematic Reviews*, (1). <https://doi.org/10.1002/14651858.MR000006.pub3>

- Hume, D. (1739). *David Hume: A Treatise of Human Nature (Second Edition)*. Oxford University Press. <https://doi.org/10.1093/actrade/9780198245872.book.1>
- Hunter, J. P., Saratzis, A., Sutton, A. J., Boucher, R. H., Sayers, R. D., & Bown, M. J. (2014). In meta-analyses of proportion studies, funnel plots were found to be an inaccurate method of assessing publication bias. *Journal of Clinical Epidemiology*, 67(8), 897–903. <https://doi.org/10.1016/j.jclinepi.2014.03.003>
- Huygens, C. (1629-1695), & van Schooten, F. (1615-1660). (1657). *De ratiociniis in ludo alae: In: Exercitationum mathematicarum liber primus [- quintus]*. Lugdunum Batavorum (Leiden, The Netherlands): ex officina Johannis Elsevirii. <https://doi.org/10.3931/e-rara-8813>
- Isserlis, L. (1918). On the Value of a Mean as Calculated from a Sample. *Journal of the Royal Statistical Society*, 81(1), 75. <https://doi.org/10.2307/2340569>
- Joober, R., Schmitz, N., Annable, L., & Boksa, P. (2012). Publication bias: What are the challenges and can they be overcome? *Journal of Psychiatry & Neuroscience*, 37(3), 149–152. <https://doi.org/10.1503/jpn.120065>
- Jovanovic, B. D., & Levy, P. S. (1997). A Look at the Rule of Three. *The American Statistician*, 51(2), 137–139.
- Kaplan, R. M., Chambers, D. A., & Glasgow, R. E. (2014). Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias. *Clinical and Translational Science*, 7(4), 342–346. <https://doi.org/10.1111/cts.12178>
- Kleijnen, J., & Knipschild, P. (1992). Review articles and publication bias. *Arzneimittelforschung*, 45(2), 587–591.
- Kolmogoroff, A. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-49888-6>
- LaPlace, Pierre Simon de. (1812). *Théorie analytique des probabilités*. Paris (France): Courcier. <https://doi.org/10.3931/e-rara-9457>
- Lau, J., Ioannidis, J. P. A., Terrin, N., Schmid, C. H., & Olkin, I. (2006). The case of the misleading funnel plot. *BMJ*, 333(7568), 597–600. <https://doi.org/10.1136/bmj.333.7568.597>
- Light, R. J., & Pillemer, D. B. (1984). *Summing up. The science of reviewing research*. Cambridge, Mass., USA: Harvard University Press.
- Louis, T. A. (1981). Confidence Intervals for a Binomial Parameter after Observing No Successes. *The American Statistician*, 35(3), 154–154.
- Lyapunov, A. M. (1901). Nouvelle forme du théorème sur la limite de probabilité. *Mémoires de l'Académie Impériale Des Sciences de St.-Petersbourg. Série VIIIe. Classe Physico-Mathématique*, 12, 1–24.
- Moher, D., Dulberg, C. S., & Wells, G. A. (1994). Statistical Power, Sample Size, and Their Reporting in Randomized Controlled Trials. *JAMA: The Journal of the American Medical Association*, 272(2), 122. <https://doi.org/10.1001/jama.1994.03520020048013>
- Moivre, A. (1733). *Approximatio ad summam terminorum binomii (a+b)n in seriem expansi*. London: Privately (Publisher not identified). Retrieved from <https://www.york.ac.uk/depts/math/histstat/demoivre.pdf>
- Moivre, A. de [1667-1754]. (1718). *The Doctrine of Chances or a Method of Calculating the Probability of Events in Play*. London: printed by W. Pearson for the author. <https://doi.org/10.3931/e-rara-10420>
- Mosteller, F. (1968). Association and Estimation in Contingency Tables. *Journal of the American Statistical Association*, 63(321), 1. <https://doi.org/10.2307/2283825>
- Murad, M. H., Chu, H., Lin, L., & Wang, Z. (2018). The effect of publication bias magnitude and direction on the certainty in evidence. *BMJ Evidence-Based Medicine*, 23(3), 84–86. <https://doi.org/10.1136/bmjebm-2018-110891>
- Pagano, M., & Gauvreau, K. (2018). *Principles of Biostatistics* (2nd ed.). Milton: CRC Press. Retrieved from <https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=5301930>
- Panagiotakos, D. B. (2008). The Value of p-Value in Biomedical Research. *The Open Cardiovascular Medicine Journal*, 2, 97–99. <https://doi.org/10.2174/1874192400802010097>
- Pearson, K., & Heron, D. (1913). On Theories of Association. *Biometrika*, 9(1–2), 159–315. <https://doi.org/10.1093/biomet/9.1-2.159>

- Pearson, Karl. (1899). XV. On certain properties of the hypergeometrical series, and on the fitting of such series to observation polygons in the theory of chance. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 47(285), 236–246. <https://doi.org/10.1080/14786449908621253>
- Pearson, Karl. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175. <https://doi.org/10.1080/14786440009463897>
- Poisson, S. D. (1837). *Recherches sur la Probabilité des jugements en matière criminelle et en matière civile, précédées des règles générales du calcul des probabilités*. Paris, France: Bachelier. Retrieved from https://www-liphy.ujf-grenoble.fr/pagesperso/bahram/Phys_Stat/Biblio/Poisson_Proba_1838.pdf
- Pólya, G. (1920). Über den zentralen Grenzwertsatz der Wahrscheinlichkeitsrechnung und das Momentenproblem: [In English: On the central limit theorem of probability calculation and the problem of moments]. *Mathematische Zeitschrift*, 8(3–4), 171–181. <https://doi.org/10.1007/BF01206525>
- Quine, M. P., & Robinson, J. (1985). Efficiencies of Chi-Square and Likelihood Ratio Goodness-of-Fit Tests. *The Annals of Statistics*, 13(2), 727–742. <https://doi.org/10.1214/aos/1176349550>
- Rumke, C. L. (1975). Implications of the Statement: No Side Effects Were Observed. *The New England Journal of Medicine*, 292(7), 372–373.
- Sachs, L. (1992). *Angewandte Statistik*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Scheid, H. (1992). *Wahrscheinlichkeitsrechnung* (Vol. 6). Mannheim: BI-Wiss.-Verl.
- Schervish, M. J. (1996). P Values: What They are and What They are Not. *The American Statistician*, 50(3), 203–206. <https://doi.org/10.1080/00031305.1996.10474380>
- Tchébychef, P. L. (1867). Des valeurs moyennes. *Journal de Mathématiques Pures et Appliquées*, 2(12), 177–184.
- von Leibniz, G. W. F. (1765). *Oeuvres philosophiques latines & françoises de feu Mr. de Leibniz*. Amsterdam (NL): Chez Jean Schreuder. Retrieved from <http://www.archive.org/stream/oeuvresphilosoph00leibuoft#page/n9/mode/2up>
- Warrens, M. J. (2008). On Association Coefficients for 2x2 Tables and Properties That Do Not Depend on the Marginal Distributions. *Psychometrika*, 73(4), 777–789. <https://doi.org/10.1007/s11336-008-9070-3>
- Wertheimer, R. (1968). Conditions. *Journal of Philosophy*, 65, 355–364.
- Yates, F. (1934). Contingency Tables Involving Small Numbers and the χ^2 Test. *The Journal of the Royal Statistical Society (Supplement)*, 1(2), 217–235.
- Yule, G. U. (1900). On the Association of Attributes in Statistics: With Illustrations from the Material of the Childhood Society, &c. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 194(252–261), 257–319. <https://doi.org/10.1098/rsta.1900.0019>
- Zwetsloot, P.-P., Van Der Naald, M., Sena, E. S., Howells, D. W., IntHout, J., De Groot, J. A., ... Wever, K. E. (2017). Standardized mean differences cause funnel plot distortion in publication bias assessments. *ELife*, 6. <https://doi.org/10.7554/eLife.24260>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).