

5-1-2002


# An Error In Statistical Logic In The Application Of Genetic Paternity Testing

Ernest P. Chiodo  
*Wayne State University*

Joseph L. Musial  
*Henry Ford Health System, Detroit, MI*

J. Sia Robinson  
*East Side Academy, Detroit, MI*

Follow this and additional works at: <http://digitalcommons.wayne.edu/jmasm>

 Part of the [Applied Statistics Commons](#), [Social and Behavioral Sciences Commons](#), and the [Statistical Theory Commons](#)

## Recommended Citation

Chiodo, Ernest P.; Musial, Joseph L.; and Robinson, J. Sia (2002) "An Error In Statistical Logic In The Application Of Genetic Paternity Testing," *Journal of Modern Applied Statistical Methods*: Vol. 1 : Iss. 1 , Article 16.

DOI: 10.22237/jmasm/1020255420

Available at: <http://digitalcommons.wayne.edu/jmasm/vol1/iss1/16>

This Regular Article is brought to you for free and open access by the Open Access Journals at DigitalCommons@WayneState. It has been accepted for inclusion in Journal of Modern Applied Statistical Methods by an authorized editor of DigitalCommons@WayneState.

## An Error In Statistical Logic In The Application Of Genetic Paternity Testing

Ernest P. Chiodo

Internal Medicine, School of Medicine  
Wayne State University

Joseph L. Musial

Department of Internal Medicine  
Henry Ford Health System

J. Sia Robinson

East Side Academy, Detroit, Michigan

---

A Bayes probability computer program was written in Fortran to examine issues related to genetic paternity testing. An application was given to demonstrate the effects improper assumptions of prior probability of culpability. The seriousness of such errors include the potential of assigning paternity to wrongly accused men, or wrongly refuting paternity.

Key words: Bayes, Paternity testing.

---

### Introduction

Genetic testing has been widely used in criminal cases as well as in cases involving establishment of paternity. In the United States, the incidence of paternity cases appears to be very large especially when compared to live birth rates. For example, the Centers for Disease Control reports that there were 3,959,417 live births nationally during calendar year 1999 (Ventura, Martin, Curtin, Menacker, & Hamilton, 2001). Among these births, a total of 1,308,560 (33%) were delivered by unmarried women. Unpublished data from The Wayne County Circuit Court, located in Southeastern Michigan, notes that there were 11,104 case filings during the year 2000 (Wayne County Circuit Court, 2001). Case filings are generated when legal paternity has not been established.

Every legal practitioner dealing with genetic testing in either a criminal or paternity context should know how this testing can lead to incorrect conclusions. The statistical assumptions made during paternity testing can cause the results of testing to be misleading and unreliable. Although the focus of this article is on a serious error in statistical methodology frequently occurring in paternity testing, the same error may also occur in criminal DNA

testing with dire consequences. An attorney practicing criminal or family law needs to understand the statistical assumptions that may cause the results of genetic testing to be misleading and unreliable. This understanding allows the knowledgeable advocate an opportunity to dispute testing results that are commonly and wrongly assumed to be infallible.

The mathematics in this article is limited in scope to the application of a simple formula. In addition, there is incentive to engage in the minimal mental effort needed to understand this article, because it provides the thoughtful attorney with a powerful advocacy tool. The central issue is the common error of assuming equal prior (pre-test) probabilities for an event in the face of ignorance concerning the actual probabilities. This common error called the "principle of indifference" may cause genetic testing using Bayes formula to be misleading (Issac, 1995).

The probability of the event A is written as  $P(A)$  and operates under the condition:  $0 \leq P(A) \leq 1.0$ . An impossible event has a probability of zero and a certain event has a probability of 1.0. Probabilities are mutually exclusive, which means there is no overlap. Tossing a coin illustrates this concept. There can be only one possible outcome: heads or tails. However, rarely are events dichotomous, straightforward probabilities. Instead, there are frequently a significant number of previous research findings with different probability levels.

This provides the entry point of Bayesian statistics. Thomas Bayes was an 18<sup>th</sup> century English clergyman who devised a formula to generate a conditional probability (Borowski & Borwein, 1991; Freund, 1973). The basic tenet of Bayesian statistics is the inclusion of conditional or prior probabilities. Often, the prior probabilities are not mechanically or deterministically generated, but rather, are based on expert judgment. Motulsky (1995) noted "usually the prior probability is not a real probability but is rather a subjective feeling. Some statisticians (Bayesians) think it is okay to convert feelings to numbers ("99% sure"

---

Ernest P. Chiodo is a physician and an attorney. He is board certified in internal medicine, occupational medicine, public health, and general preventive medicine. Contact him at 25656 Schoenherr, Suite B, Warren, MI 48089, E-mail: [echiodo@pol.net](mailto:echiodo@pol.net). Joseph L. Musial, Ph. D., is an Education Specialist for the Department of Internal Medicine at Henry Ford Health System, Detroit, MI. J. Sia Robinson is an instructor at East Side Academy, Detroit Public Schools, and adjunct lecturer, College of Education, Wayne State University. The authors acknowledge Jerome J. Fekin, Jr., for providing the case filing data, and Nicol R. Shamey, Plymouth-Canton, MI, School District.

or “70% sure”), which they define as the prior probability. Other statisticians (frequentists) think that you should never equate subjective feelings with probabilities” (p. 145).

The field of inferential statistics maintains a certain element of inaccuracy. Specifically, there are two types of errors associated with hypothesis testing: (1) a Type I error, or “false-positive”, which occurs when a researcher asserts that there was a significant finding when in fact none existed; and, (2) a Type II error, or “false negative”, which occurs when a researcher fails to observe a significant difference.

The following medical example illustrates the inherent problems associated with conditional probabilities. Assume that a 40 year-old female, who resides in an upper middle class suburb, presents with a newly diagnosed case of lung cancer to her primary care physician. The physician may use conditional probabilities generated by the Centers for Disease Control in order to determine the national incidence of this type of lung cancer among 40 year-old females. This approach appears reasonable, but caution is in order. During the history and physical, the patient may have failed to report that her former residence of many years was located near a toxic waste dump. This means that the history of environmental exposure may have contributed to the presenting cancer growth. By not including this in the construction of the prior probability, the physician may fail to consider other comorbid diseases associated with environmental exposure.

The problems associated with Bayesian statistics are also relevant to paternity cases. Suppose a man is accused of being the father of a child. He is found to have a genetic marker that only occurs in 1% of the male population. The child is tested and is also found to have the same genetic marker. The mother does not have the genetic marker. It is known that whenever a father has the marker it is always passed to the child. In this case the man contests paternity. Let:

A = The man is the father of the child

B = The child has the same genetic marker as the man

A' = The man is not the father of the child

P(A/B) = The probability that the man is the father of the child given that the child has the same genetic marker as the man.

P(B/A) = The probability that the child will have the same genetic marker as the man given that the man is the father.

P(B/A') = The probability that the child will have the genetic marker given that the man is not the father.

P(A) = The assumed prior probability before testing that the man is the father.

P(A') = The assumed prior probability before testing that the man is not the father.

Bayes formula is as follows:

$$P(A/B) = \frac{[P(B/A)P(A)]}{[P(B/A)P(A) + P(B/A')P(A')]}$$

In this case P(B/A) is 1 because there is a 100 % probability that the child will get the genetic marker if the man is the father\*. P(B/A') is 0.01 because the child has the same probability of having the genetic marker as the general population (one percent) if the man is not the father.

Recognize that only P(A) and P(A') need to be identified before plugging the values into Bayes formula. P(A') is simply 1 – P(A)\*\*. Therefore, all that remains is to identify P(A). P(A) is the assumed probability prior to testing that the man is the father of the child. In paternity testing this is often assumed to be 50% (.50). This assumption is made because there is a controversy concerning paternity. The mother of the child claims that the man is the father. The man claims that he is not the father. A prior (pre-test) probability of 50% is assumed as a default value for P(A).

If the above values are entered into Bayes formula the following result occurs:

$$P(A/B) = [(1)(0.5)] / [(1)(0.5) + (0.01)(0.5)] = 0.9901$$

Therefore, there is a greater than 99% probability of paternity when using a prior (pre-test) probability of 50% (P(A) = 0.5).

The *Joint AMA-ABA Guidelines* (Hummel, 1976; Kilmer, 1993) for likelihood of paternity are in Table 1.

Table 1. Paternity Guidelines.

Test Probability	Interpretation
< 80%	not useful
80% - 90%	undecided
90% - 95%	likely
95% - 99%	very likely
99.1% - 99.75%	extremely likely
99.8% - 99.9%	practically proven

In Michigan, paternity is presumed when the DNA profile determination determines a probability of paternity of 99% or higher (Hummel, 1976; Kilmer, 1993). In Michigan, blood tests for paternity are generally admissible in evidence at trial (Kilmer, 1993). Consequently, the man in the above example would be presumed under Michigan law to be the father of the child.

However, the results will change drastically if a lower prior (pre-test) probability of paternity is used. Instead of a 50% prior (pre-test) probability of paternity assume that P(A) is 0.001. This change to a low prior probability changes the results of Bayes formula.

$$P(A/B) = [(1)(0.001)] / [(1)(0.001) + (0.01)(0.999)] = 0.091$$

The change in the prior probability results in only a slightly greater than 9% probability of paternity. This would not result in a presumption of paternity and would in most cases be viewed as strong evidence against paternity.

The drastic change in probabilities that occur with a change in prior (pre-test) probability highlights a serious error in statistical methodology known as the "principle of indifference" (Isaac, 1995). The principle of indifference is the error of assuming equality when the actual probabilities are unknown. In paternity testing the prior (pre-test) probability is often assumed to be 50%. This assumption is made because the true probability of paternity is not known. The mother claims that the man is the father. The man denies paternity. Because it is not known who is telling the truth a fifty-fifty split on the prior (pre-test) probability is made. However, this assumption about the prior (pre-test) probability P(A) may cause a highly misleading result as the above example illustrates.

It is well known by statisticians that the principle of indifference is a serious methodological error. If there is no knowledge concerning the prior (pre-test) probability it is better to make no assumptions rather than to assume a 50-50 chance based on ignorance. Such an error leads to an assumption of a high probability that is transformed by the mathematics to an even higher probability (Isaac, 1995). A man who is able to present credible evidence that he never previously met a woman should not be assigned a prior (pre-test) probability of 50% of being the father of her child. A 50% prior (pre-test) probability is an arbitrarily value set at an unreasonably high level. Conversely, if a woman is able to produce credible evidence that she was alone with a man in an isolated location during the time period of conception, she is entitled to a prior (pre-test) probability of greater than 50%. This is needed because an inappropriately low prior (pre-test) probability can result in a misleadingly low test result. In both of the above cases the application of the principle of indifference lead to misleading results with tragic consequences.

Methodology

A Bayes probability computer program was written using Fortran 90. A total of 27 prior probabilities ranging from 0.001 to 0.90 were loaded into the program. A prior probability curve was plotted using the resulting probabilities. The Fortran 90 computer program appears below:

---

```

program one
implicit none
real::p(27)
real::ba,ba1,pp,ppp
integer::i
open(1,file='probs',status='new')
!p(27) is array of prior probabilities
!pp is prior probability
!ppp is resulting probability
!ba is P(B/A) value of 1
!ba1 is P(B/A') value of .01
ba=1.0
ba1=.01
p(1)=.001
p(2)=.002
p(3)=.003
p(4)=.004
p(5)=.005
p(6)=.006
p(7)=.007
p(8)=.008
p(9)=.009
p(10)=.01
p(11)=.02
p(12)=.03
p(13)=.04
p(14)=.05
p(15)=.06
p(16)=.07
p(17)=.08
p(18)=.09
p(19)=.1
p(20)=.2
p(21)=.3
p(22)=.4
p(23)=.5
p(24)=.6
p(25)=.7
p(26)=.8
p(27)=.9
!Call prior probabilities from array.
do i=1, 27
pp=p(i)
!Calculate resulting probability

```

```

ppp=((ba)*(pp))/((ba)*(pp)+(ba1)*(1-pp))
write(*,*) 'For a prior probability of:',pp
write(*,*) 'the resulting probability is:',ppp
write(*,*)
write(*,*)
!Write results to external file.
write(1,10) pp,ppp
10 format (2f8.4)
end do
stop
end program one

```

---

### Results

The probabilities ranged from 0.091 to 0.9989 and were negatively skewed. (See Figure 1 on next page.) The distribution begins to rapidly grow when a 0.2 prior probability was loaded. The probability of 0.091 suggests that the male defendant had a 9% chance that he was the father. Michigan's prior probability of 0.5 resulted in a 99% probability of paternity. However, this simulation was not based upon real world prior probabilities. Instead, all prior probabilities were arbitrarily loaded. The resulting probabilities, or establishing paternity, ranged from 9 to 99%.

### Conclusion

Genetic testing is seriously flawed when improper assumptions of prior probability of culpability are made. In the arena of paternity testing this has the great potential of assigning paternity to wrongly accused men. It also has the equally tragic potential of wrongly refuting paternity. In the arena of criminal law the same errors concerning assumptions about prior probability present the great risk of loss of life and liberty. The skillful legal advocate must know the potential of abuse of genetic testing and be prepared to expose the abuse when it occurs. The conditional probabilities employed by courts of law should be based upon objectivity rather than subjectivity.

The reader is invited to try various probabilities into the following Fortran 90 Bayes Probability program:

---

```

program two
implicit none
real::p
real::pp
write(*,*) 'Enter prior probability ( ≤3 decimals):'
read(*,*) p
pp=((1)*(p))/((1)*(p)+(.01)*(1-p))
write(*,*) 'The resulting probability is:',pp
stop
end program two

```

---

### Endnotes

\* *In probability mathematics a 100 percent probability is 1. A 50 percent probability is 0.5.*

\*\*  *$P(A')$  is the opposite of  $P(A)$ .  $P(A')$  is equal to one minus  $P(A)$  since in probability mathematics the sum of all the possibilities is one.*

### References

- Borowski, E. J., & Borwein, J. M. (1991). *The Harper Collins dictionary of mathematics*. New York: Harper Collins.
- Freund, J. E. (1973). *Introduction to probability*. New York: Dover Publications.
- Hummel. (1976). Joint AMA-ABA guidelines: Present status of serologic testing in problems of disputed parentage, *10 Fam LQ 247*.
- Isaac, R. (1995). *The pleasures of probability*. New York: Springer-Verlag.
- Kilmer, J. B. (1993). Paternity and surrogate parenting agreements. In J. A. Curtis, S. Bassett, & L. M. Collins (Eds.) *Michigan family law* (4<sup>th</sup> ed.). Ann Arbor, MI: The Institute of Continuing Legal Education. §17-14 and §17.15.
- Motulsky, H. (1995). *Intuitive biostatistics*. New York: Oxford University Press.
- Ventura, S. J., Martin, J. A., Curtin, S. C., Menacker, F., & Hamilton, B. E. (2001, April 17). *National vital statistics reports. Births: Final data for 1999, 49(1)*. Centers for Disease Control.
- Wayne County Circuit Court (2001). Unpublished Data. Detroit, Michigan.

# BAYES PRIOR PROBABILITY CURVE

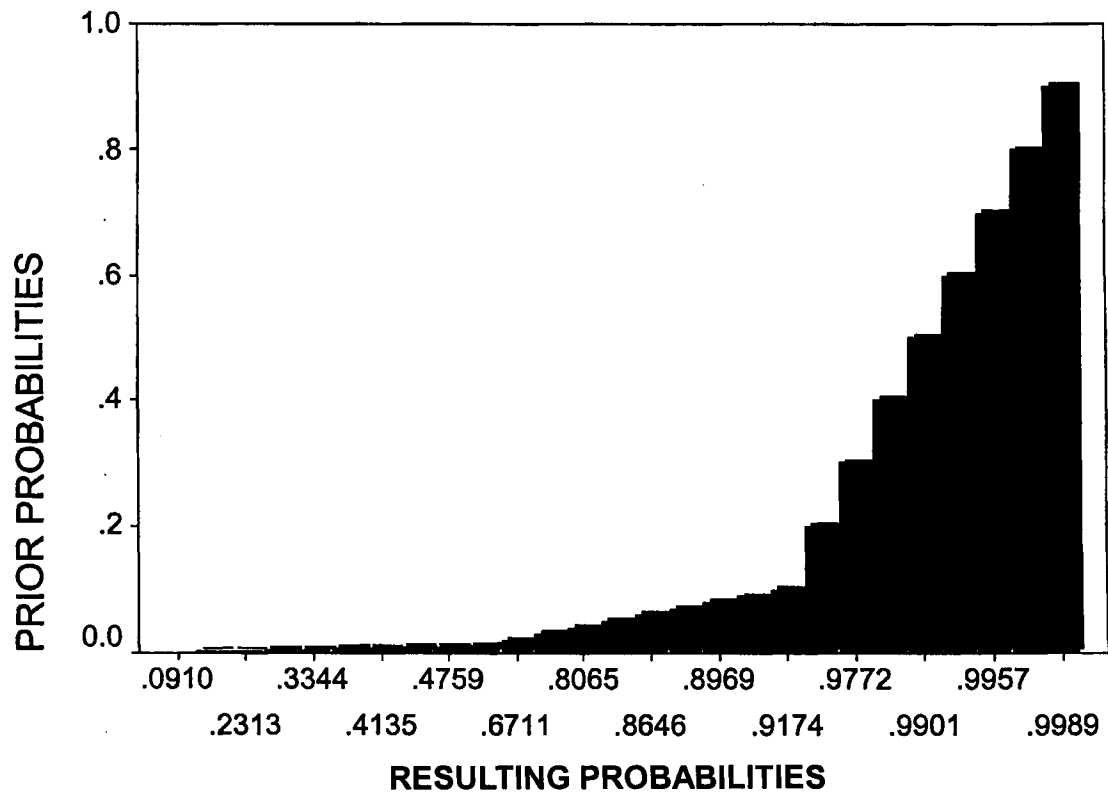


Figure 1.