

Prediction of the Probabilities of the Transmission of Genetic Traits within Bayesian Logical Inference

M. CEVRI^{a,*} AND D. ÜSTÜNDAĞ^b

^aIstanbul University, Science Faculty, Mathematics Department, Istanbul, Turkey

^bMarmara University, Arts and Science Faculty, Mathematics Department, Istanbul, Turkey

Predictions for the transmission of genetic traits along to generations are an important process for patients, their family and genetic counseling. For this purpose, Bayesian analysis in which one can include a priori knowledge taking into account all relevant information into the problem could be a useful tool to examine how disease forecasting affects its probability so that it provides a more straightforward interpretation of predictions. Therefore, we investigate here transmissions of autosomal recessive diseases along to generations within Bayesian framework. In order to do that we develop a computer code that is useful to facilitate genetic transition matrices to forecast predictions of probabilities of transmission of genetic traits by using Mathematica software, well known as an algebraic manipulation language. Furthermore, the symbolic implementation of the code is applied for the cystic fibrosis disease forecasting in humans genetics. All results show that Bayesian analysis plays a central role of prediction for probabilities of transmissions of genetic traits along generations for cystic fibrosis disease or other autosomal recessive disorders.

DOI: [10.12693/APhysPolA.130.45](https://doi.org/10.12693/APhysPolA.130.45)

PACS/topics: 02.10.Ud , 02.50.Ga, 02.50.-r , 02.70.Rr, 07.05.Kf, 42.62.Be

1. Introduction

Mendel's laws, related to the transmission of genes along generations, are actually a kind of axioms of classical probability in genetics. Therefore comprehending these probability axioms facilitates that Mendel's studies on genetics can easily be understood. By using these laws, quantifying of the risks of diseases for individuals can be done according to prior knowledge of their family pedigree and genetic process of disease being considered [1]. This was only first used in 1950's or 60's in problems of medical diagnosis and therapy planning [2] and then, in recent years, improved with use of statistical analysis based on Bayesian logical inferences, originally stated by Bayes [3–5]. It provides a more straightforward interpretation of predictions and has been employed in many genetic areas, such as classification of genotypes and estimating their relationships [6], population genetics [7, 8], gene ordering [9, 10], human-risk analysis [11, 12] and also plant disease prediction [13, 14].

The use of pedigree analysis is important for analyzing inheritance of traits in human and animal populations [15, 16]. Its goals are firstly to determine mode of inheritance such as dominant, recessive, partial dominance, sex-linked, autosomal, mitochondrial and maternal effect. Secondly, it is used to calculate the probability of an affected offspring for a given genetic cross. There are some methods that have already been used in estimating probabilities of human diseases. In this context, Boehnke and Moll [17] demonstrated that the use of maximum likelihood methods is effective for identifying

pedigrees segregating at a major locus for a quantitative trait; Lalouel and Morton [18] also introduced natural log likelihood for autosomal dominant disease and Farrer and Couples [19] suggested Bayesian probability approach to estimate probability for major gene of the Alzheimer disease in a particular family and compared it with maximum likelihood approaches. Although Ogino et al. [11, 12], taking into account various scenarios, illustrate Bayesian methods for calculating genetic risks when different mutation panels are used for the parents and the fetus we extend it for not only parents and fetus but also son, grandson and so on. On the other hand, Shoemaker et al. [20], Stansfield and Carlton [21], Carlton [22] and the others have already discussed about using classical and Bayesian probability analysis in the field of genetics, but there is a lack of work about predicting of probability of transmission of autosomal recessive traits along to generations and computational problems in real-world risk analysis associated with situations that are only hypothetical.

In this paper, we study a special kind of stochastic process called a Markov chain, where the outcome of genetic crossing of traits depends only on that of the previous one and used it to create transmission probability vector of traits in a given generation. By using this vector, we predict posterior probability of transmission of genetic traits under some assumptions. Therefore, we develop a Mathematica code that is used for predicting the probability of transmissions of cystic fibrosis (CF) disease [23–25] along to generations within Bayesian framework with Markov chain.

2. Bayesian logical inferences

Bayesian logical inference task is to infer the probability for the hypothesis H , given some data D from experi-

*corresponding author; e-mail: mcevri@gmail.com

ment and capturing also all relevant information \mathbf{I} . This can be done within the setting of Bayes' theorem which states

$$p(H|\mathbf{D}, \mathbf{I}) = \frac{p(\mathbf{D}|\mathbf{I})p(H|\mathbf{I})}{p(\mathbf{D}|\mathbf{I})}. \quad (1)$$

In this equation, the quantity $p(\mathbf{D}|\mathbf{I})$ is a global likelihood for entire class of H or evidence given some information \mathbf{I} . The quantity $p(\mathbf{D}|\mathbf{I})$ is called likelihood of H , which measures or determines a probability of observations \mathbf{D} , or the statistic under the hypothesis being tested. The quantity $p(H|\mathbf{I})$ is known as a prior probability distribution function (PDF) of H in the absence of \mathbf{D} and the quantity $p(H|\mathbf{D}, \mathbf{I})$ is a posterior PDF of H , which is a compromise between the prior information and the data. More details and references about Bayesian approach can be found in papers [26–30].

3. Pedigree analyses

Pedigree helps us infer whether a trait from a single gene is dominant or recessive so that it is used to study inheritance of a visible trait that shows differences in individuals, such as hairstyle, tongue rolling, blood type, fur, skin or hair color. Pedigree diagrams shown in Figs. 1–2 indicate effects of hereditary characteristics of living organisms among generations represented by the symbol \mathbf{G}_i whose rows indicate generations and columns single-locus ordered genotypes of offsprings, defined by $\mathbf{G} = (G_{ij})$, ($i = 1, 2, \dots, n+1, j = 1, 2$) [2, 3, 31, 32].

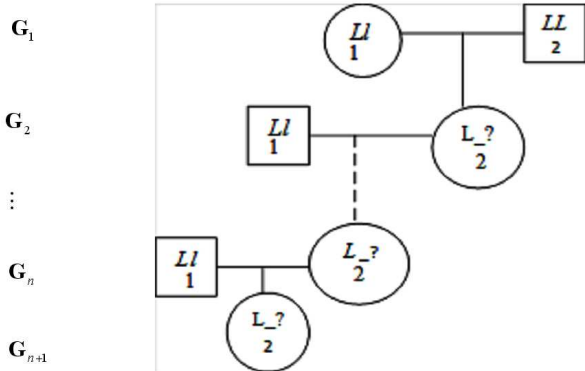


Fig. 1. Incomplete genetic pedigree for monohybrid crosses of $G_{(n+1)2}$.

In genetic science, combinations of genes on chromosomes of reproductive cells are known as “genetic crosses”, defined as a conscious breeding of two different individuals resulting in offspring that carries a portion of genetic material of both parents. Hence, let us suppose that a single gene controls the hair length of a hamster. Then, a short hair is governed by a dominant gene and represented by L while a long hair by a recessive gene and represented by l . Therefore, a hamster will be phenotypically short haired unless its genotype is (ll). Figure 1 shows a possible genetic pedigree which contains squares

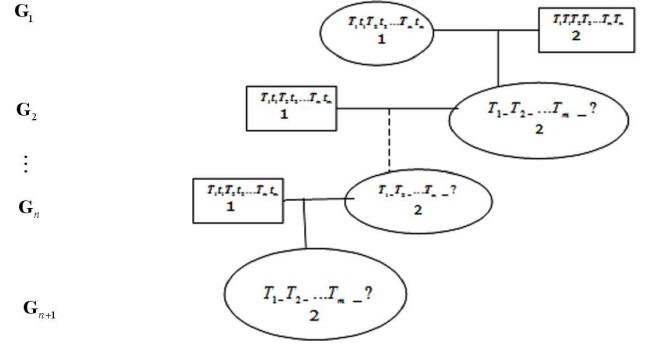


Fig. 2. Incomplete genetic pedigrees for multi-hybrid crosses of $G_{(n+1)2}$.

and circles denoted as a male and female, respectively. At the first generation \mathbf{G}_1 , both of parents are known to be phenotypically short haired and G_{11} is being genetically heterozygous (Ll). Genotypes of offsprings (G_{i2}) are short haired, but their genotypes are incompletely known ($L?$).

Given that G_{i1} 's are carriers and $G_{(n+1)2}$ has a short hair, we would like to ask what is the probability that $G_{(n+1)2}$ is being genotypically heterozygous Ll ? In order to answer this question, let us define $H = \{h_i\}_{i=1}^k$, where k is a number of hypotheses, E and I represent hypotheses of offspring's genotypes $\{LL, Ll, ll\}$, certain evidence and background information or genetic context, respectively. In this respect, our context I that represents the genotypes of parents G_{11} and G_{12} are initially known certainly as to be heterozygous short haired (Ll) and homozygous short haired (LL); our evidence E is that G_{i2} ($i = 2, \dots$) is short haired and our hypothesis h_2 is that G_{i2} 's are heterozygous short haired (Ll). Then Eq. (1) turns into the following form:

$$\begin{aligned} p(h_2|E, I) &= \frac{p(h_2|I)p(E|h_2, I)}{p(E|I)} \\ &= \frac{p(h_2|I)p(E|h_2, I)}{\sum_{i=1}^m p(h_i|I)p(E|h_i, I)}. \end{aligned} \quad (2)$$

In order to calculate the probabilities in Eq. (2) we use matrices whose components indicate probabilities of genetic crosses of traits. In this context, a relationship between the probabilities of the offspring's genotypes and that of its parents can be modeled using transition matrices used in the Markov chains. Therefore, the transition matrix \mathbf{A} whose components are the probability values of outcomes of genotypes by crossing G_{11} with LL , Ll , and ll is given in Table I.

On the other hand, let \mathbf{x} represent the prior probabilities of hypothesis $\{h_i\}_{i=1}^k$ of genotypes $\{LL, Ll, ll\}$, so that initially, $\mathbf{x}_0 = [1, 0, 0]^T$ because G_{12} is given as LL genotype in Fig. 1. After one generation,

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 = \left[\frac{1}{2}, \frac{1}{2}, 0 \right]^T, \quad (3)$$

after two generations,

TABLE I

Probabilities of genotypes obtained by crossing G_{11} with LL , Ll and ll respectively.

	$Ll \times LL$	$Ll \times Ll$	$Ll \times ll$
$p(LL)$	1/2	1/4	0
$p(Ll)$	1/2	1/2	1/2
$p(ll)$	0	1/4	1/2

$$\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 = \mathbf{A}^2\mathbf{x}_0 = \left[\frac{3}{8}, \frac{1}{2}, \frac{1}{8} \right]^T \quad (4)$$

and so on for any n ($n \in \mathbb{Z}^+$) generations we obtain

$$\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0 = \left[\frac{1}{4} + \frac{1}{2^{n+1}}, \frac{1}{2}, \frac{1}{4} - \frac{1}{2^{n+1}} \right]^T. \quad (5)$$

In the limiting case, \mathbf{x}_n implies that when a population of hamster is crossed with a heterozygous hamster, the ratio of genotypes of offsprings becomes 1:2:1. If our hypothesis h_2 is correct, $G_{(n+1)2}$ is automatically short haired, because it is dominant so that the likelihood of hypothesis $\mathbf{d} = \{p(E|h_i, I)\}_{i=1}^k$ would be $p(E|h_1, I) = p(E|h_2, I) = 1$ and $p(E|h_3, I) = 0$, respectively. By using Eq. (2) the normalization constant $p(E|I) = \mathbf{x}_n \cdot \mathbf{d}$ and the prior probability of h_2 , $p(h_2|I) = 1/2$, then the posterior PDF of h_2 given in Eq. (2) takes the following form:

$$p(h_2|E, I) = \frac{2^n}{3 \times 2^{n-1} + 1}. \quad (6)$$

It can be seen that $p(h_2|E, I)$ approaches to $2/3$ as $n \rightarrow \infty$ and imply that under assumptions about parents and all crossings with heterozygous Ll an individual born after n -th generation is likely to be heterozygous shortly haired with a probability of two-thirds. In Eq. (5), it is required to calculate high powers of matrix \mathbf{A} . Therefore, we can use similarity transformation to diagonalise the matrix \mathbf{A} and then take high powers of \mathbf{A} easily [33]:

$$\mathbf{x}_n = \mathbf{A}^n\mathbf{x}_0 = \mathbf{P}\mathbf{\Delta}^n\mathbf{P}^{-1}\mathbf{x}_0, \quad (7)$$

where $\mathbf{\Delta}$ is a diagonal matrix whose diagonal entries are eigenvalues of \mathbf{A} and \mathbf{P} is a matrix whose columns are linearly independent eigenvectors of \mathbf{A} corresponding to its eigenvalues.

Let us assume that we have m different traits denoted as $\{T_1, T_2, \dots, T_m\}$ shown in Fig. 2.

Following previous pedigree results, obtaining the probability of their transmissions of m different traits after n generation in Bayesian framework suffer from computational difficulties of \mathbf{A}^n because the dimensions of \mathbf{A} ($3^m \times 3^m$) becomes large. For example, if $m = 2$ we get

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \mathbf{0} \\ \mathbf{A}_1 & \mathbf{A}_1 & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{A}_1 \end{bmatrix}, \quad (8)$$

where

$$\mathbf{A}_1 = \begin{bmatrix} 1/4 & 1/8 & 0 \\ 1/4 & 1/4 & 1/4 \\ 0 & 1/8 & 1/4 \end{bmatrix}, \quad \mathbf{A}_2 = \begin{bmatrix} 1/8 & 1/16 & 0 \\ 1/8 & 1/8 & 1/8 \\ 0 & 1/16 & 1/8 \end{bmatrix}$$

and $\mathbf{0}$ is a (3×3) dimensional zero matrix. In order to overcome this problem we develop a symbolic code in Mathematica that has the flexibility and may easily cover any dimension of transition matrices. Either using this code or mathematical induction [34] we obtain,

$$p(h_2|E, I) = \left(\frac{2^n}{3} \right)^m \frac{1}{(2^{n-1} + 3^{-1})^m}, \quad (9)$$

under the previous assumptions [35].

4. Simulations

In previous sections we describe the use of transition matrices for determining various genetic inheritance probabilities for a simple dominant trait of hamsters, but this is a trivial application. Extension of this study in any real application could be more interesting. In this context, it will be started to predicting cystic fibrosis (CF) disease for evaluating its performance on real-world data in Bayesian framework. CF [36–38] is one of the most common, inherited, single-gene disorders in populations white Caucasian descent, such as those of Europe, North America and Australasia [37]. It is caused by mutations in CF transmembrane conductance regulator (CFTR) gene found in cells that line the lungs, digestive tract, sweat glands, and genitourinary system [36, 37]. In that disease, the mucus lining the respiratory and gastrointestinal tracts are extremely thick so that it makes individuals with this disease fight infections or absorb nutrients. CF patients are expected to die within the first years of life. Their life expectancy has lengthened with advances in diagnosis and treatment and is currently about 38 years for them [39].

Family history of disease.

TABLE II

Person	Mutat.	Genotype
grandfather (father side)	null	unaff., not carr. (CC)
grandmother (father side)	$\delta f508$	unaff., carr. (Cc)
grandfather (mother side)	null	unaff., not carr. (CC)
grandmother (mother side)	$g542x$	unaff., carr. (Cc)
father	$\delta f508$	unaff., carr. (Cc)
mother	$g542x$	unaff., carr. (Cc)
aunt (mother sister)	$g542x$	unaff., carr. (Cc)
uncle	null	unaff., not carr. (CC)
Ahmet	CF dis.	affected (cc)

On the basis of above information, we consider here proving a versatility of our proposed method through application to various common genetic counseling scenarios. Therefore, a family history of CF patient summarized in Table II is obtained by using genetic carrier test [11, 12, 40] which is a blood test and determines whether or not a carrier of the defective gene that causes CF and provided by Division of Pediatric Pulmonology of Marmara University. Let C represent the allele which codes for absence of the CF disease and c its existence. The vector of prior probabilities of hypothesis

TABLE III

Genetic testing information depends on ethnicity for cystic fibrosis disease.

Ethnicity	Carrier frequency	CF carrier testing detection rate
Caucasian	1/25	90%
Ashkenazi Jewish	1/25	97%
African American	1/65	69%
Hispanic	1/46	57%
Asian	1/94	49%

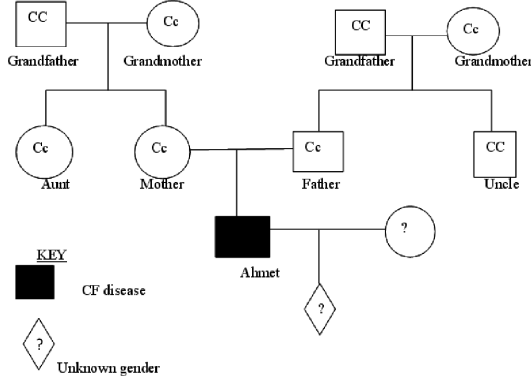


Fig. 3. A pedigree of a family with individual affected with cystic fibrosis disease.

$H = \{H_i\}_{i=1}^3$ where H_2 represents offspring's being a carrier CF but not affected (CC), H_1 represents offsprings being a carrier CF (Cc) and H_3 represents affected CF (cc) in a human population.

The CF disease is caused by homozygous mutations, such as $\Delta F508$ and G542X in the CFTR gene [40, 41] given in Table II. However, the most common mutation among them is the one denoted as $\Delta F508$ whose frequency varies with different ethnic groups (it accounts approximately for 70% of Northern European ancestry [40]). A pedigree diagram including all family members is drawn from Table II and illustrated in Fig. 3. Given the genetic context that Ahmet is a CF patient and all crossings with him and next generations after him are assumed to be Caucasian that is not affected by phenotypical CF disease, what are the probabilities that his child, grandson and so on are being genetically either heterozygous or homozygous?

If both parents are carriers, then each of their children has a 25% chance of being disorder. Therefore it is said that the probability of Ahmet being CF disease is 25%. Let us reconsider crossing with CC, Cc and either CC or Cc as mentioned previously. Since a crossing with cc contradicts our hypothesis, it is removed. In the case of absence of a family history the husband's prior carrier risk is approximately 1/25 that is given in Table III since he is a Caucasian. In a similar way, the transition matrix A is recalculated so that the frequency values of n -th generations x_n are obtained and shown in Table IV. It is seen that any person of family crossed with CC, Cc and either Cc or CC results CF disease disappears gradually as n increases. Let us consider another hypothesis, namely H_2 , that G_{i1} is unaffected, but not a CF carrier (CC). In this case, taking values of x_n in Table IV and putting into Eq. (2) values of $p(H_j|E, I)$, ($j = 1, 2$) are obtained and shown in Table V.

TABLE IV

The frequencies versus generations in different crosses with genotypes of offsprings belonging to Ahmet.

Generations	x_n		
n	Crossing with CC	Crossing with Cc	Crossing with CC or Cc
1	$[0, 0.24, 0]^T$	$[0, 0.005, 0.005]^T$	$[0, 0.1225, 0.0025]^T$
2	$[0.0576, 0.0576, 0]^T$	$[0.000025, 0.000075, 0.00005]^T$	$[0.0150, 0.0156, 0.0003]^T$
3	$[0.0276, 0.0138, 0]^T$	$[5 \times 10^{-7}, 1.25 \times 10^{-6}, 6.25 \times 10^{-7}]^T$	$[0.0037, 0.0020, 0.00003]^T$
	\vdots	\vdots	\vdots
20	$[7.64 \times 10^{-12}, 4.02 \times 10^{-13}, 0]^T$	$[5.54 \times 10^{-38}, 1.10 \times 10^{-37}, 5.54 \times 10^{-38}]^T$	$[3.56 \times 10^{-17}, 5.87 \times 10^{-18}, 1.05 \times 10^{-19}]^T$

In case of three crossings, $p(H_1|E, I)$ would decrease as n increases and asymptotically approaches 0, 0.66, and 0.14, respectively. It can clearly be seen in Fig. 4. As expected, the values of $p(H_1|E, I)$ obtained by crossing with either CC or Cc lie between those obtained by crossing with CC and Cc, respectively. However, $p(H_2|E, I)$ would increase as n increases and asymptotically approaches 1, 0.33 and 0.86. It is illustrated in Fig. 4. Moreover, it is seen that posterior PDFs of hypotheses

have a threshold about 5th generation at all crossings and reach equilibrium level after it and then carry on the same level. Although the posterior PDF of hypothesis H_1 approaches zero at crossing with CC since c gene would decrease as generation number increases, the posterior PDF of hypothesis H_2 approaches one at crossing with CC since C gene would increase as generation number increases. On the basis of the hypothesis H_1 it is said that the probability of next generation with a carrier

TABLE V

Posterior PDFs of two hypotheses for different crosses with genotypes of offsprings belonging to Ahmet.

Generations	Crossing with CC		Crossing with Cc		Crossing with CC or Cc	
	$p(H_1 E, I)$	$p(H_2 E, I)$	$p(H_1 E, I)$	$p(H_2 E, I)$	$p(H_1 E, I)$	$p(H_2 E, I)$
1	1	0	1	0	1	0
2	0.5	0.5	0.75	0.25	0.51	0.49
3	0.33	0.66	0.69	0.307	0.35	0.649
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
20	0.05	0.95	0.66	0.33	0.1413	0.858

of CF disease would approach zero but, the probability of next generation with unaffected, not a CF carrier, would reach to one, on the basis of the hypothesis H_2 .

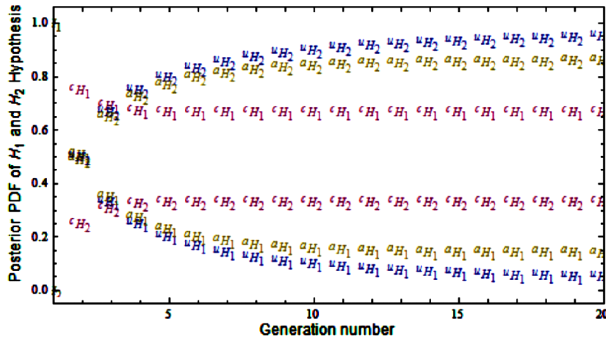


Fig. 4. Posterior PDFs of hypothesis H_1 and H_2 for Table V. Generation number: uH_2 — crossing with CC, uH_1 — crossing with CC, cH_2 — crossing with Cc, cH_1 — crossing with Cc, aH_2 — crossing with CC or Cc, aH_1 — crossing with CC or Cc.

5. Conclusions

In this study, we consider prediction of probability of transmission of genetic traits from generation to generation within a Bayesian framework. In this context, we provide a theoretical analysis of genetic crosses of breeding individuals in general and develop a symbolic Mathematica code that offers a promising of a systematic way for calculating posterior probability of hypothesis about genetic crossing of traits. The results obtained from the trivial experiment are used to predict the trend of transmission of autosomal recessive disease in human population, especially CF disease, using family information and genetic test results. Under assumptions that the genetic context that any person is a CF patient and all crossings with him/her and next generations after him/her are not affected by phenotypically CF disease, the probabilities that his child, grandson and so on are being genetically either heterozygous or homozygous reach equilibrium level and then carry on the same level. Note also that the prior and conditional probabilities often depend on ethnicity. All of the results indicates that Bayesian analysis plays a central role in predicting probability of transmissions of CF or other autosomal recessive disorders along

to generations as accurate as possible, using all available information from pedigree and/or from genetic testing. Moreover, the symbolic code in Mathematica helps us to extend to any size of transmission matrices and to calculate posterior probabilities so that it is applicable to many common scenarios generating not only a fetus but, also offsprings along generations for crossing with their different genotypes. Therefore, it deserves further investigating genetic risk assessment of autosomal recessive diseases in a variety of common clinical scenarios.

Acknowledgments

This work is a part of the project, whose name is “Bayesian Predictions of Genetic Traits” with a number 54922, supported by Scientific Research Projects Coordination Unit of Istanbul University, Istanbul, Turkey. We would like to thank Professors Bulent Karadag and Yasemin Gökdemir in the Division of Pediatric Pulmonology of Marmara University for providing CF data.

References

- [1] K. Lange, *Mathematical and Statistical Methods for Genetic Analysis*, Springer-Verlag, New York 2002.
- [2] P. Szolovits, S.P. Pauker, in: *MEDINFO 92: Proc. Seventh Conf. on Medical Informatics*, Eds. K.C. Lun, P. Degoulet, T.E. Piemme, O. Rienhoff, Elsevier, Holland 1992, p. 679.
- [3] T. Bayes, *Philos. Trans.* **53**, 370 (1763).
- [4] N. Yi, D. Zhi, *Genet. Epidemiol.* **35**, 57 (2011).
- [5] J. Hallander, P. Waldmann, P. Wang, M.J. Sillanpää, *Genetics* **185**, 645 (2010).
- [6] H. Alexander, N.A. Slade, R. Gomulkiewicz, *Theor. Appl. Genet.* **91**, 1284 (1995).
- [7] J. Zhu, J. Liu, C. Lawrence, *Bioinformatics* **14**, 25 (1998).
- [8] E. Gunel, S. Weardon, *Theor. Appl. Genet.* **91**, 534 (1995).
- [9] A. Rogatko, S. Zacks, *Am. J. Hum. Genet.* **52**, 947 (1993).
- [10] I. Hoeschele, P.M. VanRaden, *Theor. Appl. Genet.* **85**, 946 (1993).
- [11] S. Ogino, R.B. Wilson, W.W. Grody, *J. Med. Genet.* **41**, 1 (2004).
- [12] S. Ogino, R.B. Wilson, *J. Mol. Diagnost.* **6**, 1 (2004).

- [13] G.A. Carlson, *Am. J. Agricult. Econ.* **52**, 216 (1970).
- [14] J.E. Yuen, G. Hughes, *Plant Pathology* **51**, 407 (2002).
- [15] B. Kirkpatrick, in: *Proc. Int. Symp. on Bioinformatics Research and Applications (ISBRA)*, Eds.: L. Bleris, I. Măndoiu, R. Schwartz, J. Wang, Springer Verlag, USA 2012, p. 139.
- [16] A.B. Côté, B. Kirkpatrick, *NIPS: Neural Inform. Process. Syst.* **25**, 2906 (2012).
- [17] M. Boehnke, P.P. Moll, *Am. J. Hum. Genet.* **44**, 216 (1989).
- [18] J.M. Lalouel, N.E. Morton, *Hum. Hered.* **31**, 312 (1981).
- [19] L.A. Farrer, L.A. Cupples, *Am. J. Hum. Genet.* **54**, 374 (1994).
- [20] J. Shoemaker, I.S. Painter, B.S. Weir, *Trends Genet.* **15**, 354 (1999).
- [21] W. Stansfield, M.A. Carlton, *Am. Biol. Teacher* **66**, 177 (2003).
- [22] M.A. Carlton, *J. Statist. Educat.* **13**, (2005).
- [23] C.S. Richards, L.A. Bradley, J. Amos, B. Alitto, W.W. Grody, A. Maddalena, M.J. McGinnis, T.W. Prior, B.W. Popovich, M.S. Watson, G.E. Palomaki, *Genet. Med.* **4**, 379 (2002).
- [24] M.R. Knowles, M. Drumm, *Cold Spring Harb. Perspect. Med.* **2**, a009548 (2012).
- [25] M.L. Drumm, A.G. Ziady, P.B. Davis, *Annu. Rev. Pathol.* **7**, 267 (2012).
- [26] A. Gelman, J.B. Carlin, H.S. Stern, D.B. Rubin, *Bayesian Data Analysis*, Chapman and Hall, New York 1995, Ch. 7.
- [27] M. Botje, *Introduction to Bayesian Inference*, NIKHEF, Amsterdam 2009.
- [28] M. Cevri, D. Ustundag, *IET Signal Process* **6**, 673 (2012).
- [29] M. Cevri, D. Ustundag, in: *Computational Problems in Engineering*, Eds. N. Mastorakis, V. Mladenov, Springer Verlag, Bern 2014, p. 13.
- [30] D.A. Berry, *Statistics: A Bayesian Perspective*, Wadsworth Publ., Belmont, CA 1996.
- [31] W.R. Uhlmann, J.L. Schuette, B.M. Yashar, *A Guide to Genetic Counseling*, Wiley-Blackwell, 2009, p. 108.
- [32] D.L. Harti, E.W. Jones, *Genetics: Analysis of Genes and Genomes*, 6th ed., Jones and Bartlett, Sudbury, MA 2005.
- [33] D.C. Lay, *Linear Algebra and Its Applications*, 4th ed., Pearson Education, 2012.
- [34] H.R. Kenneth, *Discrete Mathematics and Its Applications*, McGraw-Hill, 2012.
- [35] M. Cevri, D. Ustundag, Y. Gökdemir, B.T. Karadag, *Selçuk J. Appl. Math.*, in press, 2016.
- [36] P. Miller, A. Hamosh, M. Macek, P. Greenberger, J. Maclean, S. Walden, R. Slavin, G. Cutting, *Am. J. Hum. Genet.* **59**, 45 (1996).
- [37] World Health Organization, *The Molecular Genetic Epidemiology of Cystic Fibrosis*, WHO, Italy 2004.
- [38] A. Anand, E. Tullis, A. Stephenson, P. Abhyankar, *J. Cyst. Fibros.* **13**, 306 (2014).
- [39] C.F. Foundation, *Cystic Fibrosis Foundation Patient Registry, 2008 Annual Data Report*, Bethesda, MD 2009.
- [40] M.F. Wildhagen, L.P. ten Kate, J.F. Habbema, *Brit. Med. Bull.* **54**, 857 (1998).
- [41] M.S. Irish, J.M. Ragi, H. Karamanoukian, D.S. Borowitz, D. Schmidt, P.L. Glick, *Pediat. Surg. Int.* **12**, 434 (1997).