American Journal of Mathematics and Statistics 2013, 3(5): 288-295
DOI: 10.5923/j.ajms.20130305.06

Estimation of Allele and Extinction Probability from Generations of an Ancestor

Babulal Seal

Department of Statistics, Burdwan University, Burdwan, W.B., India

Abstract In population genetics estimation of alleles for a gene is important to understand the character of the population. Several works are available in the literature and these are based on gene-type data of a population. They collect data from several units. But they do not collect data of an unit from the generations for estimating this. Considering the variation of genes from the generations of an unit, an estimation procedure of di-allele is attempted. Though it started from an unit, this method can be extended for many units also. However, variations over generations are considered and with this kind of data-format the extinction probability is estimated which is logical as variations of gene along generations are important. These estimates of di-alleles have been computed. Also the performance of estimators based on three generations and four generations are compared. As expected from common sense, it is seen in findings also that if one takes data from more generations then estimator becomes improved. For comparing these risks using three generations and four generations were computed and plotted. Along this line the work can be generalized by considering many units and more generations.

Keywords Branching Process, Probability Generating Function, m. l. e.

1. Introduction

Extinction probability[1] of a gene and its estimation[2] is important to understand whether a character will be vanished in the long run or not. Being motivated while working with estimation of extinction probability, it is found that the technique may be used for estimating allele also. For this we may use observations from a population or may use data from the generation chain of a single unit.

In population genetics estimation of alleles for a gene is important to understand the character of the population. Several works are available in the literature eg[3] these are based on gene-type data of the population. They collect data from several units. But they do not use data of an unit from the generations for estimating this. Considering the variation of genes from the generations of an unit, an estimation procedure of di-allele is attempted. In fact this is a kind of branching process[4]&[5]. Estimation of some branching processes are discussed in[6] &[7]. In section 2, estimation from three generation data is given by using m.l.e. Here estimation of θ, proportion of allele 'a' is given and $1-\theta$ is proportion of allele 'A'.

For simplicity we may pretend that a unit gives birth at most two offspring having $1-\theta$ = probability of being male. In that section using one unit, the estimates are given and estimation of extinction probability is also given. The risk function[8] is also calculated. In section 3 same things are carried out using four generation data. This is for simple type of branching process as described. There are many types of branching processes also e.g.[9].

In section 4 two risk functions are obtained from results of section 2 and section 3 graphically by using R- program and using squared error loss function. It is found that the second one is better for most of θ values. This is logical because more generations give the true reflection of the allele as was partially thought. It is believed that if we use more generations then estimate of θ value will improve. However, that is a point of more investigations. But up to these generations it seems to be consistent.

The reason that first one is slightly better for $\theta \leq 0.2$ may be the following. Small θ implies more 'males' but using more generations implies variation among data set i.e., mixing of 'male', 'female' information contradict that truly there are only 'males'.

2. Method for Estimation of Probability of Extinction

Suppose we want to make inference about branching process using data over generations. Let the generations be $G_0, G_1, G_2, \ldots, G_N$. We consider the simple case where a particle gives birth at most 2 offspring as

Observed Value	:	0	1	2
Probability Distribution	:	θ^2	$2\theta(1-\theta)$	$(1-\theta)^2$,$0 < \theta \leq 1$

* Corresponding author:
babulal_seal@yahoo.com (Babulal Seal)
Published online at http://journal.sapub.org/ajms

Now at G_0 if one stars with one unit, then in the following all possible paths through the three generations may be observed. In the following, possible sample values together with transition probabilities inside the 1st bracket are given for each generation (column-wise).

$$
\begin{array}{ccccc}
G_0 & & G_1 & & G_2 \\
1 \longrightarrow & \left\{\begin{array}{l}
0(p_0) \\
1(p_1) \longrightarrow \left\{\begin{array}{l} 0(p_0) \\ 1(p_1) \\ 2(p_2) \end{array}\right. \\
2(p_2) \longrightarrow \left\{\begin{array}{l} 0(p_0^2) \\ 1(2p_0p_1) \\ 2(p_0^2+2p_0p_2) \\ 3(2p_1p_2) \\ 4(p_2^2) \end{array}\right.
\end{array}\right.
\end{array}
$$

It is to be noted that the probability generating function (p.g.f.) from a single unit is

$P(s) = p_0 + p_1 s + p_2 s^2$. The p.g.f. obtain from two units X, Y will be

$$
\begin{aligned}
P_{X+Y}(s) &= (P(s))^2 = (p_0 + p_1 s + p_2 s^2)^2 \\
&= p_0^2 + p_1^2 s^2 + p_2^2 s^4 + 2p_0p_1 s + 2p_0p_2 s^2 + 2p_1p_2 s^3 \\
&= p_0^2 + 2p_0p_1 s + (p_1^2 + 2p_0p_2)s^2 + 2p_1p_2 s^3 + p_2^2 s^4
\end{aligned}
$$

So the possible paths together with probabilities will be:

Table 1. Paths and probabilities using 3 generations

Paths	Probabilities	Expression in terms of θ
$s_1 = (1,0,0)$	p_0	θ^2
$s_2 = (1,1,0)$	p_0p_1	$2\theta^3(1-\theta)$
$s_3 = (1,1,1)$	p_1^2	$4\theta^2(1-\theta)^2$
$s_4 = (1,1,2)$	p_1p_2	$2\theta(1-\theta)^3$
$s_5 = (1,2,0)$	$p_0^2p_2$	$\theta^4(1-\theta)^2$
$s_6 = (1,2,1)$	$2p_0p_1p_2$	$4\theta^3(1-\theta)^3$
$s_7 = (1,2,2)$	$p_1^2p_2 + 2p_0p_2^2$	$6\theta^2(1-\theta)^4$
$s_8 = (1,2,3)$	$2p_1p_2^2$	$4\theta(1-\theta)^5$
$s_9 = (1,2,4)$	p_2^3	$(1-\theta)^6$

Now let us suppose that we have $X_1, X_2, \dots, X_n$ i.i.d observations from the above distribution. Then the joint density becomes

$$
\begin{aligned}
&f(x_1, x_2, \dots, x_n \mid \theta) \\
&= (\theta^2)^{t_1}\left(2\theta^3(1-\theta)\right)^{t_2}(4\theta^2(1-\theta)^2)^{t_3} \\
&\quad (2\theta(1-\theta)^3)^{t_4}(\theta^4(1-\theta)^2)^{t_5}(4\theta^3(1-\theta)^3)^{t_6} \\
&\quad (6\theta^2(1-\theta)^4)^{t_7}(4\theta(1-\theta)^5)^{t_8}(1-\theta)^{t_9}
\end{aligned}
$$

where $\sum_{i=1}^{9} t_i = n$ and t_i=number of s_i sample for i=1,2,…,9

Therefore

$$
\begin{aligned}
&f(x_1, x_2, \dots, x_n; \theta) \\
&= 2^{t_2+t_4}\, 4^{t_3+t_6+t_8}\, 6^{t_7}\, \theta^{2t_1+3t_2+2t_3+t_4+4t_5+3t_6+2t_7+t_8} \\
&\qquad (1-\theta)^{t_2+2t_3+3t_4 2t_5+3t_6+4t_7+5t_8+t_9}
\end{aligned}
$$

In order to obtain maximum likelihood estimate

$$
\begin{aligned}
&\frac{\partial}{\partial\theta}\log L(x \mid \theta) = (2t_1 + 3t_2 + 2t_3 + t_4 + 4t_5 + 3t_6 + 2t_7 + t_8)\frac{1}{\theta} \\
&\quad -(2t_2 + t_3 + 3t_4 + 2t_5 + 3t_6 + 4t_7 + 5t_8 + t_9)\frac{1}{(1-\theta)} \\
&\qquad = \frac{A}{\theta} - \frac{B}{(1-\theta)}
\end{aligned}
$$

where

$$
\begin{aligned}
A &= 2t_1 + 3t_2 + 2t_3 + t_4 + 4t_5 + 3t_6 + 2t_7 + t_8 \\
\& \ B &= 2t_2 + t_3 + 3t_4 + 2t_5 + 3t_6 + 4t_7 + 5t_8 + t_9
\end{aligned}
$$

Now

$$
\begin{aligned}
\frac{\partial}{\partial\theta}\log L(x \mid \theta) = 0 &\Rightarrow \frac{A}{\theta} = \frac{B}{(1-\theta)} \\
&\Rightarrow A - A\theta = B\theta \\
&\Rightarrow \hat{\theta} = \frac{A}{A+B}
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial\theta^2}\log L(x \mid \theta) &= -(2t_1 + 3t_2 + 2t_3 + t_4 + 4t_5 + 3t_6 \\
&\quad + 2t_7 + t_8)\frac{1}{\theta^2} \\
&\quad -(2t_2 + t_3 + 3t_4 + 2t_5 + 3t_6 + 4t_7 + 5t_8 + t_9)\frac{1}{(1-\theta)^2} \\
&< 0, \forall\theta
\end{aligned}
$$

Hence above estimate is a maximum likelihood estimate.

Now based on one sample path the estimates are as follows:

Table 2. Paths, summary of statistic and estimates of parameter

Paths	Summary of statistic	Estimate of parameters
s_1	$(t_1 = 1, all\ zeros)$	$\hat{\theta} = \frac{2}{2+0} = 1$
s_2	$(t_2 = 1, all\ zeros)$	$\hat{\theta} = \frac{3}{4}$
s_3	$(t_3 = 1, all\ zeros)$	$\hat{\theta} = \frac{2}{4} = \frac{1}{2}$
s_4	$(t_4 = 1, all\ zeros)$	$\hat{\theta} = \frac{1}{4}$
s_5	$(t_5 = 1, all\ zeros)$	$\hat{\theta} = \frac{4}{6} = \frac{2}{3}$
s_6	$(t_6 = 1, all\ zeros)$	$\hat{\theta} = \frac{3}{6} = \frac{1}{2}$
s_7	$(t_7 = 1, all\ zeros)$	$\hat{\theta} = \frac{1}{3}$
s_8	$(t_8 = 1, all\ zeros)$	$\hat{\theta} = \frac{1}{6}$
s_9	$(t_9 = 1, all\ zeros)$	$\hat{\theta} = \frac{0}{1} = 0$

From this estimate one can get idea of being extinct or fate of the population.

Probability of extinction and its estimate in this case:

Now to get extinction probability we solve the following equation:

$$P(s_0) = s_0$$

i.e., $p_0 + p_1 s_0 + p_2 s_0^2 = s_0$

$$\Rightarrow p_2 s_0^2 - (1-p_1)s_0 + p_0 = 0$$

$$\Rightarrow s_0 = \frac{(1-p_1) \pm \sqrt{(1-p_1)^2 - 4p_0 p_2}}{2p_2}$$

where

$$\sqrt{(1-p_1)^2 - 4p_0 p_2}$$

$$= \sqrt{\left(1 - 2\theta(1-\theta)\right)^2 - 4\theta^2(1-\theta)^2}$$

$$\sqrt{1 - 4\theta(1-\theta) + 4\theta^2(1-\theta)^2 - 4\theta^2(1-\theta)^2}$$

$$= \sqrt{1 - 4\theta(1-\theta)}$$

$$= \sqrt{(1 - 4\theta + 4\theta^2)} = 1 - 2\theta$$

Therefore,

$$s_0 = \frac{1 - 2\theta(1-\theta) \pm (1-2\theta)}{2(1-\theta)^2}$$

$$= \frac{1 - 2\theta + 2\theta^2 + 1 - 2\theta}{2(1-\theta)^2} \; or \; \frac{1 - 2\theta + 2\theta^2 - 1 + 2\theta}{2(1-\theta)^2}$$

$$= \frac{2 - 4\theta + 2\theta^2}{2(1-\theta)^2} \; or \; \frac{2\theta^2}{2(1-\theta)^2}$$

$$= 1 \; or \; \left(\frac{\theta}{(1-\theta)}\right)^2$$

So the extinction probability is $\left(\frac{\theta}{(1-\theta)}\right)^2$ & it's estimate will be $\left(\frac{\hat{\theta}_x}{(1-\hat{\theta}_x)}\right)^2$ where $\hat{\theta}_x$ are given in (2.1).

Remark 2.1

It is to be noted that

$$\frac{\theta}{(1-\theta)} \le 1 \Rightarrow \theta \le \frac{1}{2}$$

Now let us choose θ_0 be such that $\theta \le \theta_0 < \frac{1}{2}$ then

$$\left(\frac{\theta}{(1-\theta)}\right)^2 \le \alpha^2$$

$$\Rightarrow \frac{\theta}{(1-\theta)} \le \alpha$$

$$\Rightarrow \theta \le \alpha(1-\theta)$$

$$\Rightarrow \theta + \alpha\theta \le \alpha$$

$$\Rightarrow \theta \le \frac{\alpha}{1+\alpha}$$

So if $\theta_0 = \theta_\alpha = \frac{\alpha}{1+\alpha}$ then $\forall\, \theta \le \theta_0$ we have $\left(\frac{\theta}{(1-\theta)}\right)^2 \le \alpha^2$ as $\frac{\theta}{1-\theta}$ is an $\uparrow$ function of θ

If we collect huge data and obviously this is not a difficult task, then from each subset of data one can obtain estimates. From several such estimate one can find the prior distribution of θ and if it is seen that it has negligible weights beyond $(0, \theta_\alpha)$ then there is nothing to be worried about of being extinct.

Remark 2.2

Actually θ is controlled by the sex gene factor and from one generation we can not get good information about that. For this we should read pattern of transition from one generation to another. But for practical purposes we may not get data for many generations. But at least using 3 or 4 generations we should study the transitions.

Remark 2.3

Above can be used in testing the hypothesis regarding the sex gene i.e. $H_0: \theta = \theta_0$ or one can get idea whether two communities have equal θ's through usual Neyman Pearson test. Before doing that we are to compute its statistical curvature as the above family is a curved exponential family. We can say that tests of curved family holds good if its statistical curvature is small i.e., it does not deviate from regular 2-parameter exponential family much.

3. General Case

Now depending on the lengths of chain let us see the behavior of the estimator and compare this case with previous one via MSE

If we take four generation then the sample paths, probabilities will be as in the following.

Table 3. Paths and probabilities of such paths

Paths (data)	Probabilities $[p_0 = \theta^2,\ p_1 = 2\theta(1-\theta),\ p_2 = (1-\theta)^2]$	Expression in terms of θ
$s_1 = (1,0,0,0)$	p_0	θ^2
$s_2 = (1,1,0,0)$	$p_0 p_1$	$2\theta^3(1-\theta)$
$s_3 = (1,1,1,0)$	$p_0 p_1 p_1$	$2^2\theta^4(1-\theta)^2$
$s_4 = (1,1,1,1)$	$p_1 p_1 p_1$	$2^3\theta^3(1-\theta)^3$
$s_5 = (1,1,1,2)$	$p_1 p_1 p_2$	$2^2\theta^2(1-\theta)^4$
$s_6 = (1,1,2,0)$	$p_0{}^2\, p_1 p_2$	$2\theta^5(1-\theta)^3$
$s_7 = (1,1,2,1)$	$p_1 p_2(2p_0 p_1)$	$2^3\theta^4(1-\theta)^4$
$s_8 = (1,1,2,2)$	$p_1 p_2(p_1{}^2 + 2p_0 p_2)$	$2^2 3\theta^3(1-\theta)^5$
$s_9 = (1,1,2,3)$	$p_1 p_2(2p_1 p_2)$	$2^3\theta^2(1-\theta)^6$
$s_{10} = (1,1,2,4)$	$p_1 p_2 p_2{}^2$	$2\theta(1-\theta)^7$
$s_{11} = (1,2,0,0)$	$p_0{}^2 p_2$	$\theta^4(1-\theta)^2$
$s_{12} = (1,2,1,0)$	$p_0 p_2(2p_0 p_1)$	$2^2\theta^5(1-\theta)^3$
$s_{13} = (1,2,1,1)$	$p_1 p_2(2p_0 p_1)$	$2^3\theta^4(1-\theta)^4$
$s_{14} = (1,2,1,2)$	$p_2(2p_0 p_1)p_2$	$2^2\theta^3(1-\theta)^5$
$s_{15} = (1,2,2,0)$	$p_2(p_1{}^2 + 2p_0 p_2)p_0{}^2$	$2.3\theta^6(1-\theta)^4$
$s_{16} = (1,2,2,1)$	$p_2(p_1{}^2 + 2p_0 p_2)(2p_0 p_1)$	$2^3 3\theta^5(1-\theta)^5$
$s_{17} = (1,2,2,3)$	$p_2(p_1{}^2 + 2p_0 p_2)(2p_1 p_2)$	$2^3 3\theta^3(1-\theta)^7$
$s_{18} = (1,2,2,4)$	$p_2(p_1{}^2 + 2p_0 p_2)p_2{}^2$	$2.3\theta^2(1-\theta)^8$
$s_{19} = (1,2,3,0)$	$p_2(2p_1 p_2)p_0{}^3$	$2^2\theta^7(1-\theta)^5$
$s_{20} = (1,2,3,1)$	$p_2(2p_1 p_2)(3p_1 p_0{}^2)$	$2^3 3\theta^6(1-\theta)^6$

$s_{21} = (1,2,3,2)$	$p_2(2p_1p_2)(3p_0{p_1}^2 + 3p_2{p_0}^2)$	$2^2 3.5\theta^5(1-\theta)^7$
$s_{22} = (1,2,3,3)$	$p_2(2p_1p_2)({p_1}^3 + 6p_0p_1p_2)$	$2^4 5\theta^4(1-\theta)^8$
$s_{23} = (1,2,3,4)$	$p_2(2p_1p_2)(3p_0{p_2}^2 + 3p_2{p_1}^2)$	$2^2 3.5\theta^3(1-\theta)^9$
$s_{24} = (1,2,3,5)$	$p_2(2p_1p_2)(3p_1{p_2}^2)$	$2^3 3\theta^2(1-\theta)^{10}$
$s_{25} = (1,2,3,6)$	$p_2(2p_1p_2){p_2}^3$	$2^2\theta(1-\theta)^{11}$
$s_{26} = (1,2,4,0)$	$p_2{p_2}^2{p_0}^4$	$\theta^8(1-\theta)^6$
$s_{27} = (1,2,4,1)$	$p_2\,{p_2}^2(4p_1{p_0}^3)$	$2^3\theta^7(1-\theta)^7$
$s_{28} = (1,2,4,2)$	$p_2\,{p_2}^2\left(4p_2{p_0}^3 + \frac{4!}{2!\,2!}{p_1}^2{p_0}^2\right)$	$2^3 7\theta^6(1-\theta)^8$
$s_{29} = (1,2,4,3)$	$p_2\,{p_2}^2\left(\frac{4!}{2!}p_2p_1{p_0}^2 + 4{p_1}^3p_0\right)$	$56\,\theta^5(1-\theta)^9$
$s_{30} = (1,2,4,4)$	$p_2\,{p_2}^2\left(\frac{4!}{2!\,2!}{p_2}^2{p_0}^2 + \frac{4!}{2!}p_2{p_1}^2p_0 + {p_1}^4\right)$	$70\,\theta^4(1-\theta)^{10}$
$s_{31} = (1,2,4,5)$	$p_2\,{p_2}^2\left(4p_2{p_1}^3 + \frac{4!}{2!}p_0p_1{p_2}^2\right)$	$56\,\theta^3(1-\theta)^{11}$
$s_{32} = (1,2,4,6)$	$p_2\,{p_2}^2\left(4p_0{p_2}^3 + \frac{4!}{2!\,2!}{p_1}^2{p_2}^2\right)$	$16\,\theta^2(1-\theta)^{12}$
$s_{33} = (1,2,4,7)$	$p_2\,{p_2}^2(4p_1{p_2}^3)$	$2^3\theta(1-\theta)^{13}$
$s_{34} = (1,2,4,8)$	$p_2\,{p_2}^2{p_2}^4 = {p_2}^7$	$(1-\theta)^{14}$

So $L(\theta) = constant.\,\theta^A(1-\theta)^B$

$\log L(\theta) = constant + A\,log\theta + B\log\ (1-\theta)$

$$\frac{\partial log L(\theta)}{\partial\theta} = \frac{A}{\theta} + \frac{B}{1-\theta} = 0$$

$$\Rightarrow \frac{A}{B} = \frac{\theta}{1-\theta}$$

$$\Rightarrow \hat{\theta} = \frac{A}{A+B}$$

Estimator of the parameter using above paths

Table 4. Estimator of the parameter

Paths (data)	Probability $\hat{\theta}$
s_1	1
s_2	$\frac{3}{4}$
s_3	$\frac{2}{3}$
s_4	$\frac{1}{2}$
s_5	$\frac{1}{3}$
s_6	$\frac{5}{8}$
s_7	$\frac{1}{2}$
s_8	$\frac{3}{8}$
s_9	$\frac{1}{4}$
s_{10}	$\frac{1}{8}$
s_{11}	$\frac{2}{3}$
s_{12}	$\frac{5}{8}$
s_{13}	$\frac{1}{2}$
s_{14}	$\frac{3}{8}$
s_{15}	$\frac{3}{5}$
s_{16}	$\frac{1}{2}$
s_{17}	$\frac{3}{10}$
s_{18}	$\frac{1}{5}$
s_{19}	$\frac{7}{12}$
s_{20}	$\frac{1}{2}$
s_{21}	$\frac{5}{12}$
s_{22}	$\frac{1}{3}$
s_{23}	$\frac{1}{4}$
s_{24}	$\frac{1}{6}$
s_{25}	$\frac{1}{12}$
s_{26}	$\frac{4}{7}$
s_{27}	$\frac{1}{2}$
s_{28}	$\frac{3}{7}$
s_{29}	$\frac{5}{14}$
s_{30}	$\frac{2}{7}$
s_{31}	$\frac{3}{14}$
s_{32}	$\frac{5}{7}$
s_{33}	$\frac{1}{14}$
s_{34}	0

4. Comparison of Two Estimators Using 3 Generations and 4 Generations

R-codes for computations of the risk functions and plotting of the risk functions.

```
s1<-function(x)
{
sum({(1-x)^2}*x^2,{(3/4-x)^2}*2*(1-x)*x^3,
{(1/2-x)^2}*4*(1-x)^2*x^2,{(1/4-x)^2}*2*x*(1-x)^3,
{(2/3-x)^2}*x^4*(1-x),{(1/2-x)^2}*4*x^3*(1-x)^3,{(1/3-x)^2}*6*x^2*(1-x)^4,
{(1/6-x)^2}*4*x*(1-x)^5,{(0-x)^2}*(1-x)^6
)
}
x<-c(.1,.2,.3,.4,.5,.6,.7,.8,.9,1)
r<-numeric(length(x))
for(i in 1:length(x))
{
r[i]<-s1(x[i])
}
```

```
t<-c(1, 3/4, 4/6, 1/2, 1/3, 5/8, 1/2, 3/8, 1/4, 1/8, 2/3, 5/8, 1/2, 3/8, 3/5, 1/2, 3/10, 1/5, 7/12, 1/2, 5/12,
1/3, 1/4, 1/6, 1/12, 4/7, 1/2, 3/7, 5/14, 2/7, 3/14, 5/7, 1/4, 0)
s2<-function(x)
 {
sum({((t[1]-x)^2)*x^2}, {((t[2]-x)^2)*(x^3)*2*(1-x)} ,{((t[1+2]-x)^2)*4*(x^4)*((1-x)^2)}, {((t[1+3]-x)^2)* (2^3)*(x^3)*((1-x)^3)}

,{((t[1+4]-x)^2)*(2^2)*(x^2)*((1-x)^4)},{((t[1+5]-x)^2)*(2^1)*(x^5)*((1-x)^3)},{((t[1+6]-x)^2)*(2^3)*(x^4)*((1-x)^4)
},{((t[1+7]-x)^2)*(2^2)*3*(x^3)*((1-x)^5)}

,{((t[1+8]-x)^2)*(2^3)*(x^2)*((1-x)^6)},{((t[1+9]-x)^2)*(2^1)*(x^1)*((1-x)^7)},{((t[1+10]-x)^2)*(2^0)*(x^4)*((1-x)^2
```

```
)},{((t[1+11]-x)^2)*(2^2)*(x^5)*((1-x)^3)}

,{((t[1+12]-x)^2)*(2^3)*(x^4)*((1-x)^4)},{((t[1+13]-x)^2)*(2^2)*(x^3)*((1-x)^3)},{((t[1+14]-x)^2)*(2^1)*3*(x^6)*((1
-x)^4)},{((t[1+15]-x)^2)*(2^3)*3*(x^5)*((1-x)^5)}

,{((t[1+16]-x)^2)*(2^3)*3*(x^3)*((1-x)^7)},{((t[1+17]-x)^2)*(2^1)*3*(x^2)*((1-x)^8)},{((t[1+18]-x)^2)*(2^2)*(x^7)*(
(1-x)^5)},{((t[1+19]-x)^2)*(2^3)*3*(x^6)*((1-x)^6)}

,{((t[1+20]-x)^2)*(2^2)*3*5*(x^5)*((1-x)^7)},{((t[1+21]-x)^2)*(2^4)*5*(x^4)*((1-x)^8)},{((t[1+22]-x)^2)*(2^2)*3*5*
(x^3)*((1-x)^9)},{((t[1+23]-x)^2)*(2^3)*3*(x^2)*((1-x)^10)}

,{((t[1+24]-x)^2)*(2^2)*(x^1)*((1-x)^11)},{((t[1+25]-x)^2)*(2^0)*(x^8)*((1-x)^6)},{((t[1+26]-x)^2)*(2^3)*(x^7)*((1-
x)^7)},{((t[1+27]-x)^2)*(2^2)*7*(x^6)*((1-x)^8)},{((t[1+28]-x)^2)*(2^0)*5*6*(x^5)*((1-x)^9)},{((t[1+29]-x)^2)*(2^0)*
7*(x^5)*((1-x)^10)}

,{((t[1+30]-x)^2)*(2^0)*56*(x^3)*((1-x)^11)},{((t[1+31]-x)^2)*(4^2)*(x^2)*((1-x)^12)},{((t[1+32]-x)^2)*(2^3)*(x^1)
*((1-x)^13)},{((t[1+33]-x)^2)*(2^0)*(x^0)*((1-x)^14)}

)
}
x<-c(.1,.2,.3,.4,.5,.6,.7,.8,.9,1)
r1<-numeric(length(x))
for(i in 1:length(x))
{
r1[i]<-s2(x[i])
}
plot(x,r,xlab=expression(theta),ylab="MSE / Risk Functions",main="Plot of Risk Functions of the Estimators
 of Allele Proportion Using Data over Generations")
lines(x,r,lty = 1, pch = 6,lwd=2 )
lines(x,r1,lty = 2, pch = 4,lwd = 6)
legend(.8, .08, c("Using 3 generations","Using 4 generations"),lwd=2:6,
,lty=1:1, merge = TRUE, bg = 'gray90')
```

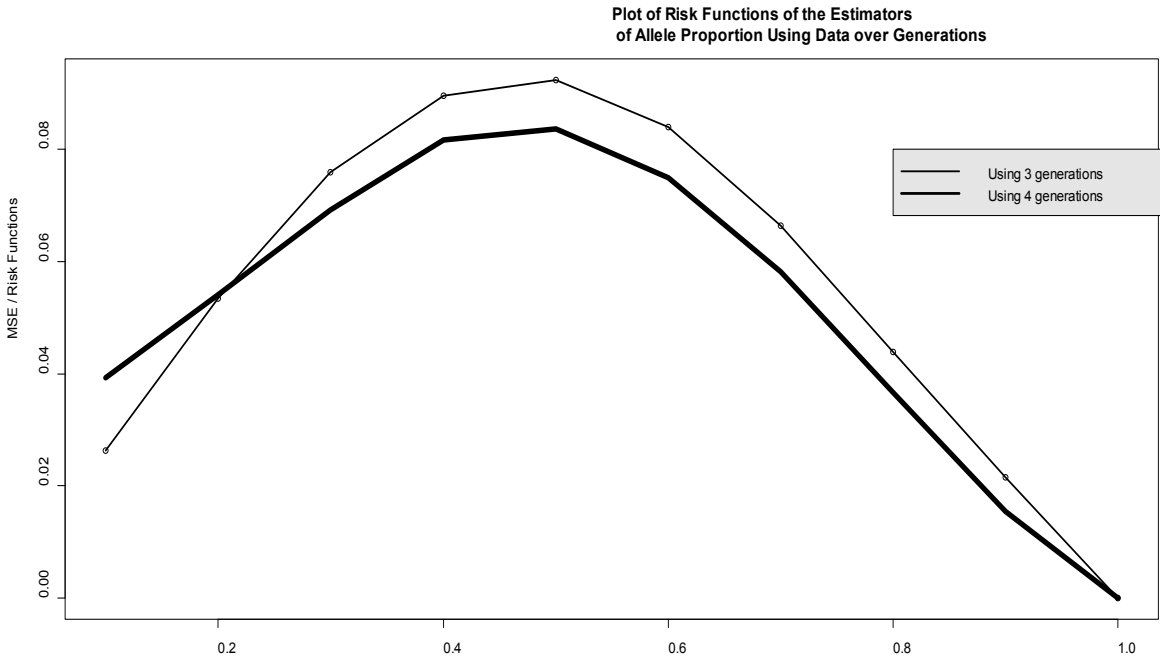

Figure 1. Comparison of two risk function

5. Concluding Remarks

There are enough scopes for elaborating the inference part like observing samples with more generations. The testing part in this regard should be carefully handled, as the family distribution here is a curved exponential family.

REFERENCES

[1] Galton, F., and Watson, H. W. (1874): On the probability of extinction of families. J. of Anthr. Inst., 6, 138-144.

[2] Guttorp, P. (1991): Statistical Inference for Branching Processes. Wiley, New York.

[3] Ben Hui Liu (1997): Statistical Genomics. CRC Press.

[4] Arthreya, K. B., and Ney, P. E. (1972): Branching Processes. Springer, New York.

[5] Mode, C. J. (1971): Multi type Branching Processes. Elsevier, New York.

[6] Keiding, N., and Lauritzen, S. (1978): Marginal Maximum Likelihood Estimates and Estimation of the Offspring's Mean in Branching Process. Scand. J. Stat., 5, 106-110.

[7] Nagaev, A. E. (1967): On Estimating the Expected Number of Direct Descendants of a Particle in a Branching Process. Theory Prob. Appl.,12, 314-320.

[8] Berger, J. O. (1985): Statistical Decision Theory and Related Topics. Springer.

[9] Seal, B. (2013): A Generalized Branching Process with Dependent Branching Times. International Journal of Statistics and Analysis. 3. 159-166