Genetic Test of Chukchi Circuit Hypothesis Based on Dirichlet Distribution and Whaling Management Based on People's Need

Master Thesis of Environment Economics

Department of Economics

University of Oslo

Master Student: YingKang Wei

Approved by: ______________________

Authorized to Offer Degree: ______________________

Date: 2007-01-10

ACKNOWLEDGMENTS

I would like to express sincere gratitude to my supervisor, Prof. Tore Schweder for his guidance and support. I thank him also for providing me an opportunity to grow as a student in the unique research environment he creates.

I would also like to express sincere thanks to Prof. Olav Bjerkholt for sound advice and kind help.

Dedication

This work is dedicated to my wife, Huaitian Bu, my son, YuHeng Wei, and my parents, especially to my mother, who is still fighting cancer in Beijing. I love them forever.

Abstraction

Bowhead whales is an endangered species but they are still hunted by Inuits. How to protect the endangered species and at the same time fulfill people's need is an important topic in economics. Resource economics and Welfare economics are two branches of economics and both of them give some answer. However, economic efficiency may not be preferred because it doesn't lead to society welfare in realities. Considering the bowhead whale problem, a simulation was done in this paper to find the catch limit which can fulfill the Inuits' need and meanwhile, protect the whale from extinction.

It is very important to make clear the structure of the stock to protect the biodiversity of the species. Chukchi Circuit hypothesis was put forward and Jorde. et. al (2004) found the "Oslo bump" of the pair-wise microsatellite differences in data from whales landed during the autumn migration at Barrow. The simulation of BCB whale migration hypothesis based on Dirichlet distribution has never been made before and will be done in this paper. 54 samples landed at Barrow in Autumn were studied, with genetic measurement on 11 loci. The results show that to get a neat bump, very extreme parameter will be required.

Index

Chapter 1　Introduction

Over the past three decades, we have become increasingly aware of environmental problems facing communities, nations, and the world. Species extinction is one of these problems. Most of the species extinctions from 1000 AD to 2000 AD are due to human activities, in particular destruction of plant and animal habitats. Some studies show that about one of eight known plant species is threatened with extinction. This figure indicates unsustainable ecological practices. Almost all scientists acknowledge that the rate of species loss is greater now than at any time in human history, with extinctions occurring at rates hundreds of times higher than background extinction rates.

Should we allow the hunting of endangered species? The answer that springs to mind is probably, "Of course not!". However, issues such as these are much more complicated in real life. The bowhead whale has been declared an endangered species, but it is also an animal traditionally hunted by the Inuit, which is not only a source of food, but also part of their culture. Should they still be allowed to hunt the whales and how far should we go to protect the animals?

The bowhead whale in the Bering-Chukchi-Beufort Seas(BCB) were killed in large numbers for their large quantities of baleen and oil. In 19th century, Victorian crinolines made from baleen plates came out of fashion and a large number of bowhead whales were killed. Fig1-1 is the number of bowhead whales killed from 1848 to 2004. In 1850 and 1852, more than 2000 Bowhead whales were killed each year. It was absolutely a kind of depredation. There was no record in 1855 and 1856. In the following 50 years since 1857, the number of the Bowhead whales killed every year has been around 300 to 500 steadily. Whaling had not been effectively stopped until 1915 due to poor catch rates and collapse of markets(Bockstoce, 1986; Bockstoce & Burns, 1993). The Bowhead whale has been officially protected since 1946 by International Whaling Committee (IWC). However, these whales are still harvested in Alaska by Inuits. Currently the stock is estimated to hold about 8,000-12,000 bowhead whales world-wide[1].

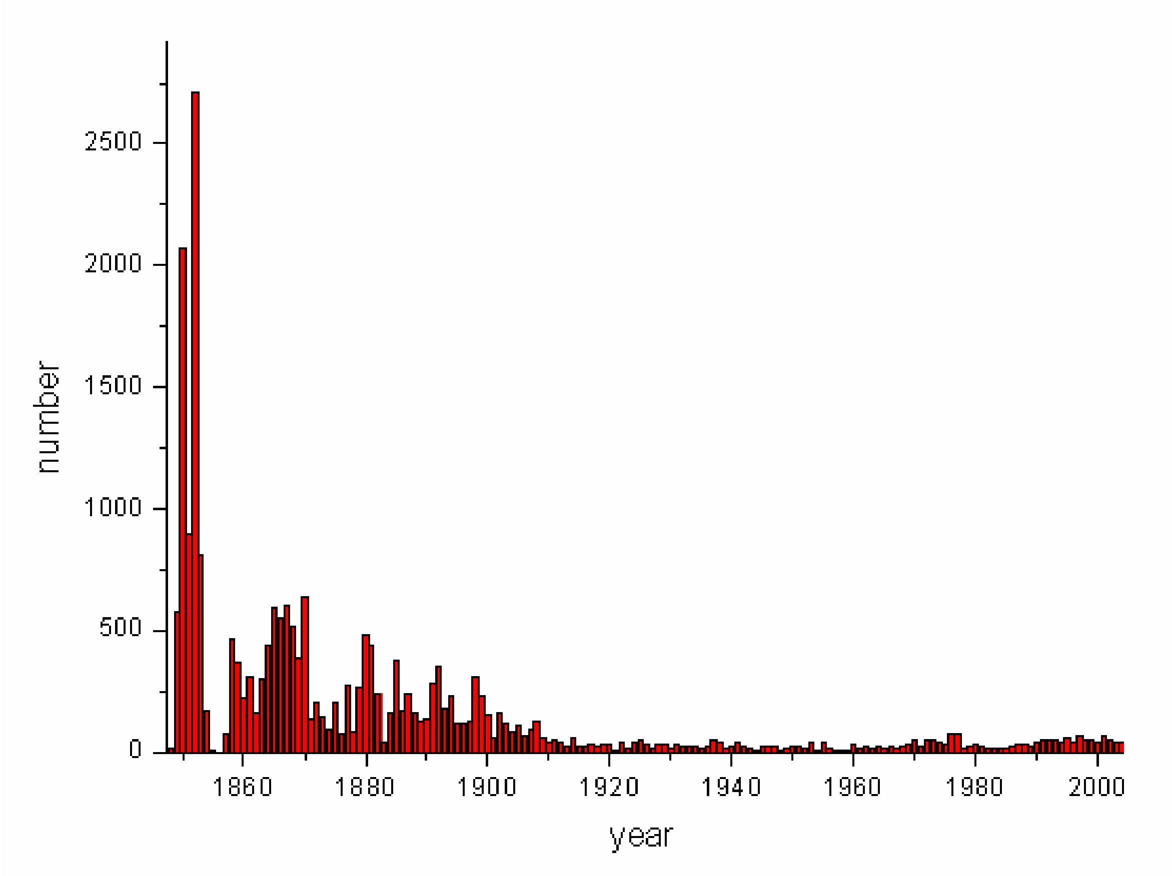


Fig1-1 BCB bowhead whale killed from 1848 to 2004 (IWC)

Many elements of uncertainty may be able to affect the whale's fertility rate, such as the feeding place, the environmental pollution, the sudden climate changes and so on. When the uncertainty is larger, the danger of extinction is larger and more conservative management is needed. The purpose of the management is to guarantee the sustainable growth of the whale's population and in the meantime meets people's needs. According to IWC reference (J. Cetacean Res. Manage.7 (suppl). 2005, pp18), the Committee has until recently been unable to provide a satisfactory management advice. A management procedure for determining catch limit has been developed by the IWC scientific committee. In October 2002, the catch limits given below were agreed for the Bering-Chukchi-Beaufort Seas stock of bowhead whales at a Special Meeting of the Commission: A total of up to 280 bowhead whales can be landed in the period 2003 - 2007, with no more than 67 whales struck in any year (and up to 15 unused strikes may be carried over each year).

A species is usually divided into many small groups, which have their own habit areas. The smaller the group, the higher risk of extinction will be with the same level of uncertainty. Due to uncertainty with respect to population structure (Jorde et al, 2004), the implementation of the management procedure will be reviewed in 2007. If the population turn out to be composed of two or more sub-stocks, the additional safeguards must be developed to ensure that no sub-population is being severely depleted in the long run.

Five different stocks are presently recognized within the species' distribution range. These stocks were all subject to extensive commercial hunting in the 19th century, and the BCB stock is the only one that is recovering successfully after commercial whaling ended in 1914 (Bockstoce, 1986; Bockstoce &Burns, 1993). BCB bowhead whales stay in arctic waters, making a northern migration in the spring to rich feeding waters. They migrate south again in the fall to breed and calve. Knowledge of its social behaviour, including the route of migration, is limited because they can be observed for only part of the year due to adverse weather conditions and ice cover in winter.

The current dominant hypothesis about the feeding migration of the BCB Bowhead whales is Baseline Hypothesis: they migrate in one group along the Alaskan coast to the Beaufort Sea in spring and back again in fall. However, according to Russian observations and oceanographic data, there might be two distinct patterns in feeding migration for whales wintering in the Bering Sea (Melnikov 2004; Bogoslovskaya 2003), which leads to Chukchi Circuit hypothesis:

There are two sub-populations, which are set to be E and W migrate in different ways:

E-bowheads migrate in spring along the Alaskan coast to the Beaufort Sea and back again in fall. They pass Barrow at each migration, and are subject to harvest there.

The W-whales leave the Beering Sea in late May and June and heads northwest on the Chukotka coast. Some summer further north and migrate south the Barrow canyon and passes Barrow on their way back to the Beering Sea in the autumn. Most of the commercial catch in 1848-1914 was also taken in central and western Bering and Chukchi Seas, which supports the existence of a western compound of the population.

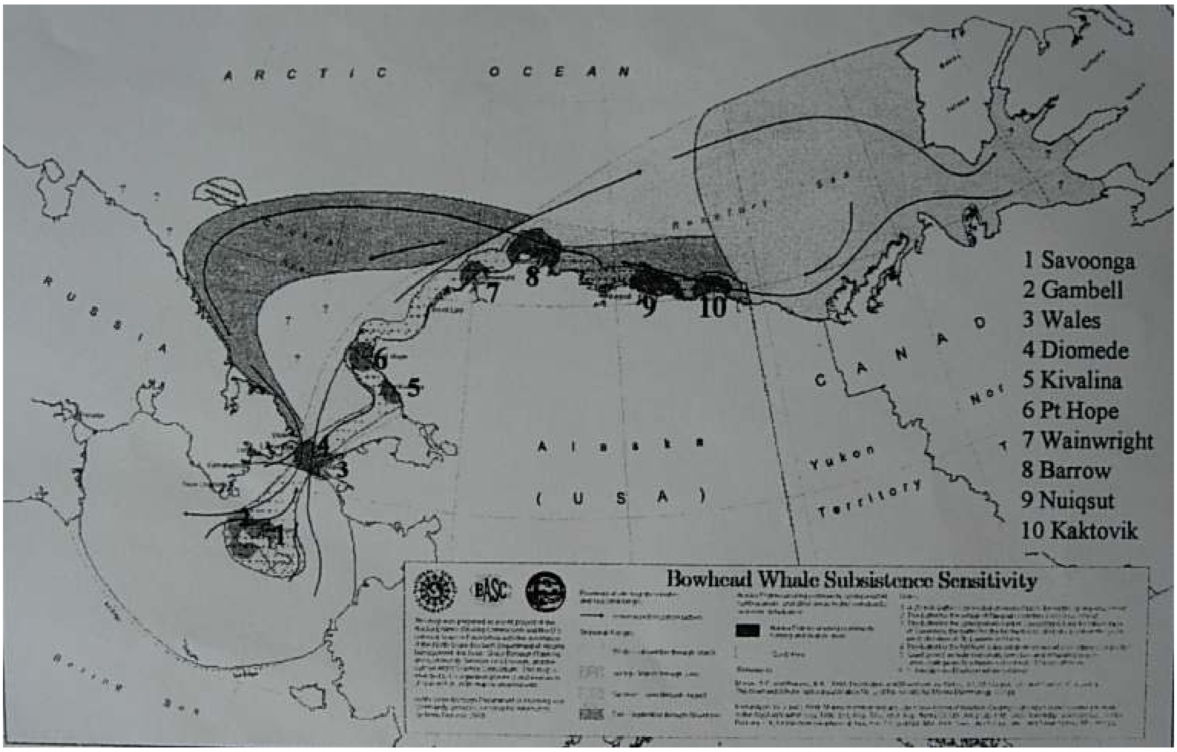


Fig 1-2 Map of the western Arctic showing the range of the BCB stock of bowhead whales.

In migration, whales are hunted for subsistence by aboriginal peoples, primarily in Alaska. Of the ten villages that hunt in Alaska, Barrow is the only community that harvests bowheads both during the spring and autumn migration, and lands about 70% of the total catch. According to traditional Inuit knowledge, both in spring and fall, bowhead whales pass Barrow in three pulses. If Chukchi Circuit hypothesis is right, then one of the fall pulses might be a W-pulse. A pattern based on pair-wise microsatellite differences is evident in data from whales landed during the autumn migration at Barrow. Larger genetic differences occurred when paired samples were 5-11 days apart than they were 0-5 or 12-25 days apart. This pattern is hardly consistent with the population being well mixed. (Jorde et al, 2004) The pattern is named “The Oslo Bump”. The Oslo bump might have resulted from temporal segregation between two population compounds at Barrow in the fall migration.

Dirichlet distribution is usually used to simulate the genetic structure. The simulation of BCB whale migration hypothesis based on Dirichlet distribution has never been made before. In this study, I have constructed a genetic model based on Dirichlet distribution, with which to simulate the distribution of microsatellite alleles in bowhead whales. The null hypothesis that there is only one migration route was tested based on the simulation result.

I will first review very briefly some basic genetics to help the reader understand how genetics data can be used to study a possible segregation in the bowhead whale population. Biodiversity is the diversity of life and is very important to our human being. I will discuss briefly the threats to biodiversity and the management to it. Then I will review the Dirichlet distribution and its special case, the Beta distribution.

In Chapter 3, I will explain how potential genetics data can be simulated by the Dirichlet distribution, and how simulation can help to investigate whether the observed pattern found by Jorde et al (2004) and called the Oslo bump is consistent with the particular hypothesis of stock structure and differential migration called the Chukchi circuit hypothesis.

In Chapter 4, I will try to discuss how to fulfil people's need, which is an important topic of economics. I will introduce the answer given by resource economics and welfare economics and put forward the difficulties in realities.

In Chapter 5, I will bring together what I found in Chapter 3-4, and give my very brief and tentative simulation and conclusions with respect to the economics of need and ecological constraints in the case of bowhead whales for Alaskan Inuits.

Chapter 2 Genes, Biodiversity and Dirichlet Distribution

Introduction to genes

A gene is a stretch of DNA (deoxyribonucleic acid) coding for a polypeptide chain. A protein is made up of one or more polypeptides. DNA is present in all cells. The genetic information in DNA is coded in the sequence of four nucleotides, abbreviated according to the identity of the nitrogenous base that each contains A, G, T or C. DNA molecules normally consist of two complementary helical strands held together by pairing between the bases: A in one strand is paired with T in another and G in one strand is paired with C in another. The entire DNA in a cell is collectively called the genome. Genome size is typically expressed as the amount of DNA in a reproductive cell (sperm or egg), and it differs greatly among species.

Genes are arranged in linear order along microscopic threadlike bodies called chromosomes. A typical chromosome contains several thousand genes. The position of a gene along a chromosome is called the locus of the gene. At each locus there may be different sequence of information which leads to the difference in every individual of the group. All these different sequence are called alleles.

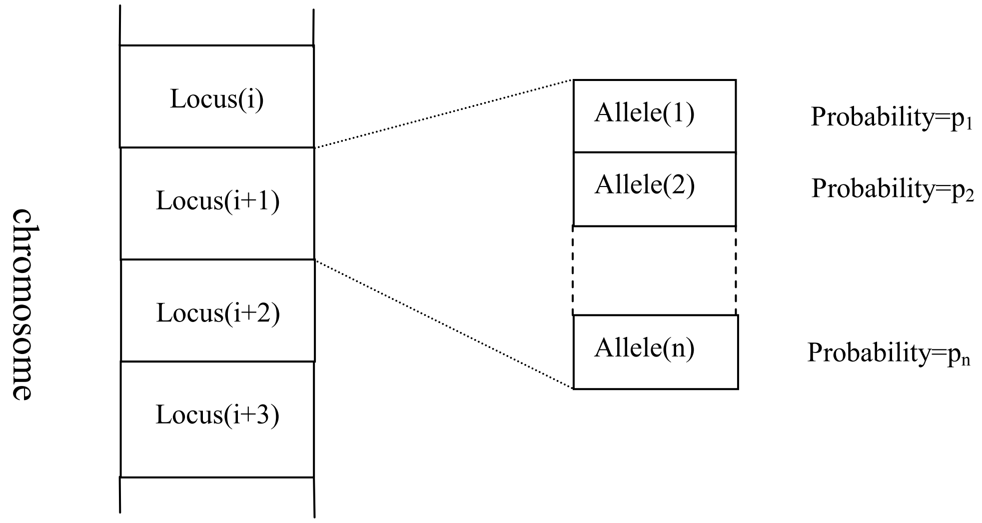


Fig.2-1 Gene Structure: chromosome, locus and allele

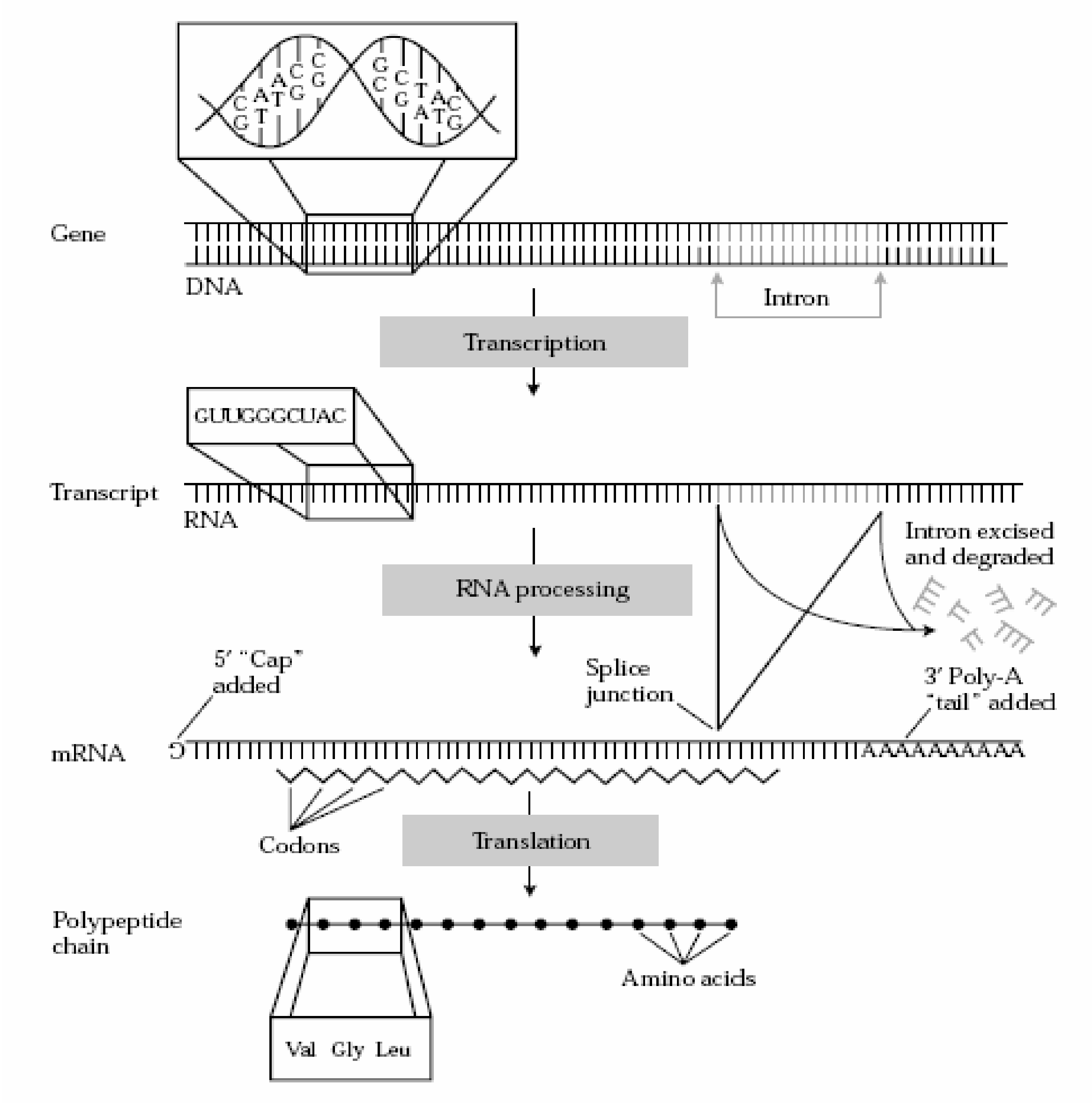


Fig.2-2 Gene Expression

The process of creating proteins from the genetic code in DNA is called gene expression, which is shown in Fig1-3. This process is done by the help of RNA (ribonucleic acid). First, the DNA untangles itself by the help of some special enzyme and copies one of its strands to RNA by the rule: A is paired with U and G is paired with C. There is U in RNA instead of T. Then RNA goes out of cell and enters cytoplasm where the protein is created. In the translated part of the messenger RNA, each adjacent group of three nucleotides constitutes a coding group, which specifies a corresponding amino acid subunit in the polypeptide chain. (Wikipedia)

Hardy-Weinberg equilibrium

Diploid means each individual cell contains two copies of each type of chromosome, one inherited from its mother through the egg and one inherited from its father through the sperm. These two copies are the alleles of the gene in that individual. If the two alleles at a locus are same (indistinguishable according to any particular experimental criterion), then the individual is homozygous at the locus under consideration. If the two alleles at a locus are distinguishable, then the individual is heterozygous at the locus. (Jeffrey K. Conner,2004)

If we mate two individuals that are heterozygous (e.g., **Bb**) for a trait, we find that

- 25% of their offspring are homozygous for the dominant allele (BB)
- 50% are heterozygous like their parents (Bb)
- 25% are homozygous for the recessive allele (bb) and thus, unlike their parents, express the recessive phenotype.

This is what Mendel found when he crossed monohybrids. It occurs because Meiosis separates the two alleles of each heterozygous parent so that 50% of the gametes will carry one allele and 50% the other. When the gametes are brought together at random, each **B**/**b** carrying egg will have a 1 in 2 probability of being fertilized by a sperm carrying **B**/**b**:

	0.5**B**	0.5**b**
0.5**B**	0.25**BB**	0.25**Bb**
0.5**b**	0.25**Bb**	0.25**bb**

If the frequency of two alleles in an **entire population** of organisms is not exactly the same, will it stay stable? Let us take as a hypothetical case, a population of hamsters in which

- 80% of all the gametes in the population carry a dominant allele for black coat (**B**)
- 20% carry the recessive allele for gray coat (**b**).

Random union of these gametes (right table) will produce a generation:

	0.8**B**	0.2**b**
0.8**B**	0.64**BB**	0.16**Bb**
0.2**b**	0.16**Bb**	0.04**bb**

Will gene b eventually disappear? No, because all the gametes (child) formed by **BB** will contain allele **B** as well and one-half the gametes formed by heterozygous (**Bb**). So, 80% (0.64 + 0.5*0.32) of the pool of gametes formed by this generation with contain **B**. All the gametes of the **bb** group (4%) will contain **b** but one-half of the gametes of the heterozygous (**Bb**) will as well. So 20% (0.04 + 0.5*0.32) of the gametes will contain **b**. The proportion of allele **b** in the population has remained the same. The heterozygous group ensures that each generation will contain 4% gray hamsters.

Now let us look at an algebraic analysis of the same problem using the expansion of the binomial $(\mathbf{p}+\mathbf{q})^2$. Since $(\mathbf{p}+\mathbf{q})^2 = \mathbf{p}^2 + \mathbf{2pq} + \mathbf{q}^2$ which indicates:

- $\mathbf{p} + \mathbf{q} = 1$
    - $\mathbf{p}^2$ = the fraction of the population homozygous for **p**
    - $\mathbf{q}^2$ = the fraction homozygous for **q**
    - **2pq** = the fraction of heterozygous

In our example, **p** = 0.8, **q** = 0.2, and thus $(0.8 + 0.2)^2 = (0.8)^2 + 2(0.8)(0.2) + (0.2)^2 = 0.64 + 0.32 + 0.04$ .The algebraic method enables us to work backward as well as forward. So the recessive genes do not tend to be lost from a population no matter how small their representation is, so long as certain conditions are met (without mutation, genetic drift, migration and natural selection). The gene frequencies and genotype ratios in a randomly-breeding population remain constant from generation to generation which is known as the **Hardy-Weinberg** law in honor of the two men who first realized the significance of the binomial expansion to population genetics and hence to evolution. (Hardy, 1908, Stern, 1943)

Biodiversity

Biodiversity is the diversity of life, which means the variety and variability of all living organisms and the ecological complexes of which they are parts. It includes genetic diversity, species diversity and ecosystem diversity (Wikipedia):

genetic diversity ------ diversity of genes within a species. There is a genetic variability among the populations and the individuals of the same species.

species diversity ------ diversity among species in an ecosystem.

ecosystem diversity --- diversity at a higher level of organization, the ecosystem.

If the gene is the fundamental unit of natural selection, according to E. O. Wilson, the real biodiversity is the genetic diversity.

The Importance of biodiversity

Biodiversity has contributed in many ways to the development of human culture, and, in turn, human communities have played a major role in shaping the diversity of nature at the genetic, species, and ecological levels. The contribution can be classified in two ways: ecological role of biodiversity and economic role of biodiversity.

All species provide at least one function in an ecosystem. Each function is an integral part of regulating the species balance, species diversity and species health: all aspects which are intrinsic for the ecosystem as a whole to survive and prosper. Ecosystems also provide various infrastructures of production (soil fertility, pollinators of plants, predators, decomposition of wastes...) and services such as purification of the air and water, stabilization and moderation of the climate, decrease of flooding, drought, and other environmental disasters [12].

Research suggests that a more diverse ecosystem is better able to withstand environmental stress and consequently is more productive. The loss of a species is thus likely to decrease the ability of the system to maintain itself or to recover from damage or disturbance [12]. Just like an ecosystem with high biodiversity, a species with high genetic diversity may have a greater chance of adapting to environmental change. In other words, the more species comprising an ecosystem, the more resilient and stable the ecosystem is likely to be.

For all humans, biodiversity is a resource for daily life. Many see biodiversity as a reservoir of resources to be drawn upon for the manufacture of food and pharmaceutical products. Since biological resources represent an ecological interest for the community, their economic value is also increasing. New products are developed because of biotechnologies, and new markets created. For society, biodiversity is also a field of activity and profit.

Finally, biodiversity is important because each species can give scientists some clue as to how life evolved and will continue to evolve on Earth. In addition, biodiversity helps scientists understand how life functions and the role of each species in sustaining ecosystems. The availability of unique genetic material for each living species may have incalculable value as evidenced by medical and genetic research that can lead to discoveries that may reduce mortality.

Threats to biodiversity

Elevated rates of extinction are being driven by human consumption of organic resources, especially related to tropical forest destruction. While most of the species that are becoming extinct are not food species, their biomass is converted into human food when their habitat is transformed into pasture, cropland, and orchards. Because an ecosystem decreases in stability as its species are made extinct, studies warn that the global ecosystem is destined for collapse if it is further reduced in complexity. Factors contributing to loss of biodiversity are: overpopulation, deforestation, pollution (air pollution, water pollution, soil contamination) and global warming or climate change, driven by human activity.

Some characterize loss of biodiversity not as ecosystem degradation but by conversion to trivial standardized ecosystems (e.g., monoculture following deforestation). In some countries lack of property rights or access regulation to biotic resources necessarily leads to biodiversity loss (degradation costs having to be supported by the community).

The widespread introduction of exotic species by humans is a potent threat to biodiversity. When exotic species are introduced to ecosystems and establish self-sustaining populations, the endemic species in that ecosystem, that have not evolved to cope with the exotic species, may not survive. The exotic organisms may be either predators, parasites, or simply aggressive species that deprive indigenous species of nutrients, water and light. These exotic or invasive species often have features due to their evolutionary background and environment that makes them very competitive, and similarly makes endemic species very defenseless and/or uncompetitive against these exotic species.

The rich diversity of unique species across many parts of the world exist only because they are separated by barriers, particularly seas and oceans, from other species of other land masses, particularly the highly fecund, ultra-competitive, generalist "super-species". These are barriers that could never be crossed by natural processes, except for many millions of years in the future through continental drift. However humans have invented ships and airplanes, and now have the power to bring into contact species that never have met in their evolutionary history, and on a time scale of days, unlike the centuries that historically have accompanied major animal migrations. As a consequence of the above, if humans continue to combine species from different eco-regions, there is the potential that the world's ecosystems will end up dominated by a very few, aggressive, cosmopolitan "super-species" [12].

Biodiversity management

The conservation of biological diversity has become a global concern. Although not everybody agrees on extent and significance of current extinction, most consider biodiversity essential. At national levels a Biodiversity Action Plan is sometimes prepared to state the protocols necessary to protect an individual species. Usually this plan also details extant data on the species and its habitat.

There are basically two main types of conservation options, in-situ conservation and ex-situ conservation. In-situ is usually seen as the ideal conservation strategy. However, its implementation is sometimes unfeasible. For example, destruction of rare or endangered species' habitats sometimes requires ex-situ conservation efforts. Furthermore, ex-situ conservation can provide a backup solution to in-situ conservation projects. Some believe both types of conservation are required to ensure proper preservation (Wikipedia). An example of an in-situ conservation effort is the setting-up of protection areas. Examples of ex-situ conservation efforts, by contrast, would be planting germplasts in seed banks. Such efforts allow the preservation of large populations of plants with minimal genetic erosion.

Introduction to Gamma distribution

A gamma distribution is a general type of statistical distribution that is related to the beta distribution. The general formula for the probability density function (pdf) of the gamma distribution is:

$$f(x) = \frac{(\frac{x-\mu}{\beta})^{\gamma-1}\exp\left(-\frac{x-\mu}{\beta}\right)}{\beta\Gamma(\gamma)} \qquad x \geq \mu; \gamma, \beta > 0 \tag{2.1}$$

where γ is the shape parameter, μ is the location parameter, β is the scale parameter, and Γ is the gamma function which has the formula:

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt \tag{2.2}$$

The case where $\mu = 0$ and $\beta = 1$ is called the standard gamma distribution. The equation for the standard gamma distribution reduces to

$$f(x) = \frac{x^{\gamma-1}e^{-x}}{\Gamma(\gamma)} \qquad x \geq 0; \gamma > 0 \tag{2.3}$$

The formula for the cumulative distribution function of the gamma distribution is

$$F(x) = \frac{\Gamma_x(\gamma)}{\Gamma(\gamma)} \qquad x \geq 0; \gamma > 0 \tag{2.4}$$

where Γ is the gamma function defined above and $\Gamma_x(a)$ is the incomplete gamma function. The incomplete gamma function has the formula

$$\Gamma_x(a) = \int_0^x t^{a-1} e^{-t} dt \tag{2.5}$$

The method of moments estimators of the gamma distribution are

$$\hat{\gamma} = (\frac{\bar{x}}{s})^2 \qquad \hat{\beta} = \frac{s^2}{\bar{x}}$$

where $\bar{x}$ and s are the sample mean and standard deviation, respectively.

The following is the plot of the gamma probability density function [2].

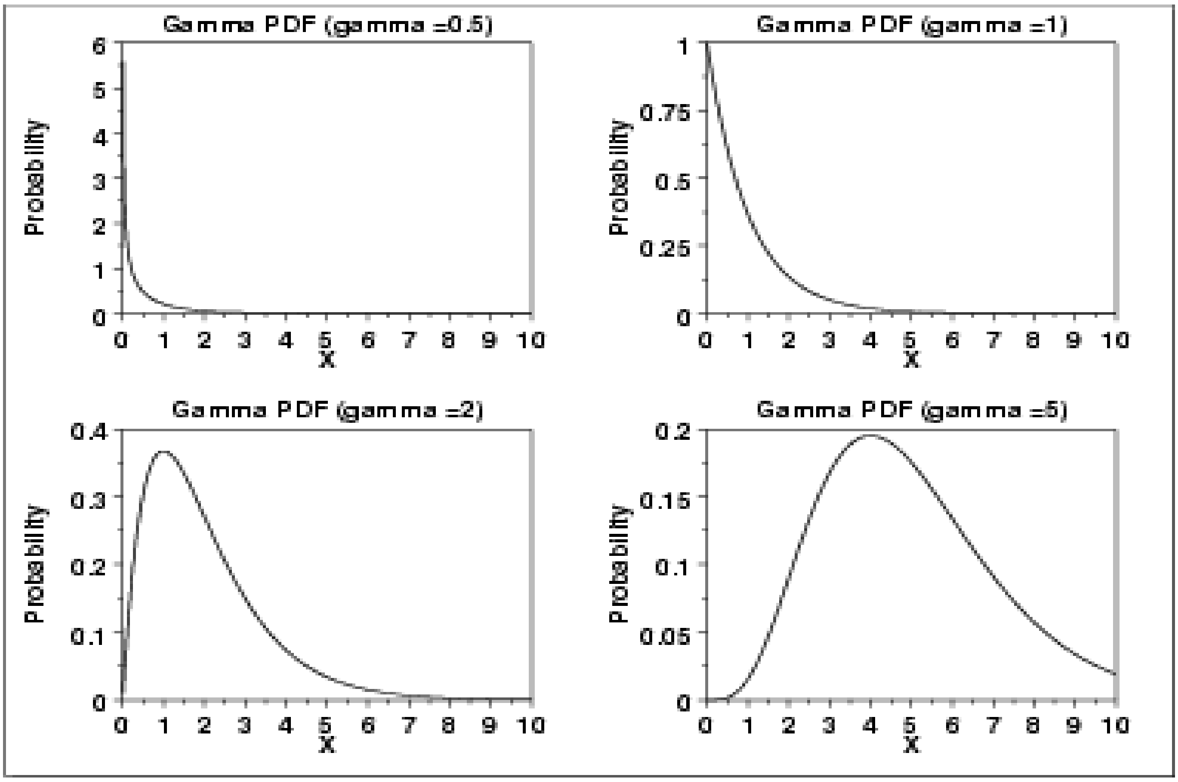


Fig. 2-3 the p.d.f of some Gamma distributions $\gamma = 0.5, 1, 2, 5$

Introduction to Beta distribution

The general formula for the p.d.f of the beta distribution is:

$$f(x) = \frac{(x-a)^{\alpha_1 - 1}(b-x)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)(b-a)^{\alpha_1 + \alpha_2 - 1}}, \qquad a \le x \le b; \ \alpha_1, \alpha_2 > 0 \qquad (2.6)$$

Where α_1 and α_2 are the shape parameters, a and b are the lower and upper bounds, respectively, of the distribution, and $B(\alpha_1, \alpha_2)$ is the beta function:

$$B(\alpha_1, \alpha_2) = \int_0^1 t^{\alpha_1 - 1}(1-t)^{\alpha_2 - 1} dt = \frac{\Gamma(\alpha_1)\Gamma(\alpha_2)}{\Gamma(\alpha_1 + \alpha_2)} \qquad (2.7)$$

The case where a = 0 and b = 1 is called the standard beta distribution. The equation for the standard beta distribution is

$$f(x) = \frac{x^{\alpha_1 - 1}(1-x)^{\alpha_2 - 1}}{B(\alpha_1, \alpha_2)}, \quad 0 \le x \le 1; \ \alpha_1, \alpha_2 > 0 \qquad (2.8)$$

The following is the plot of the beta probability density function for four different values of the shape parameters [2].

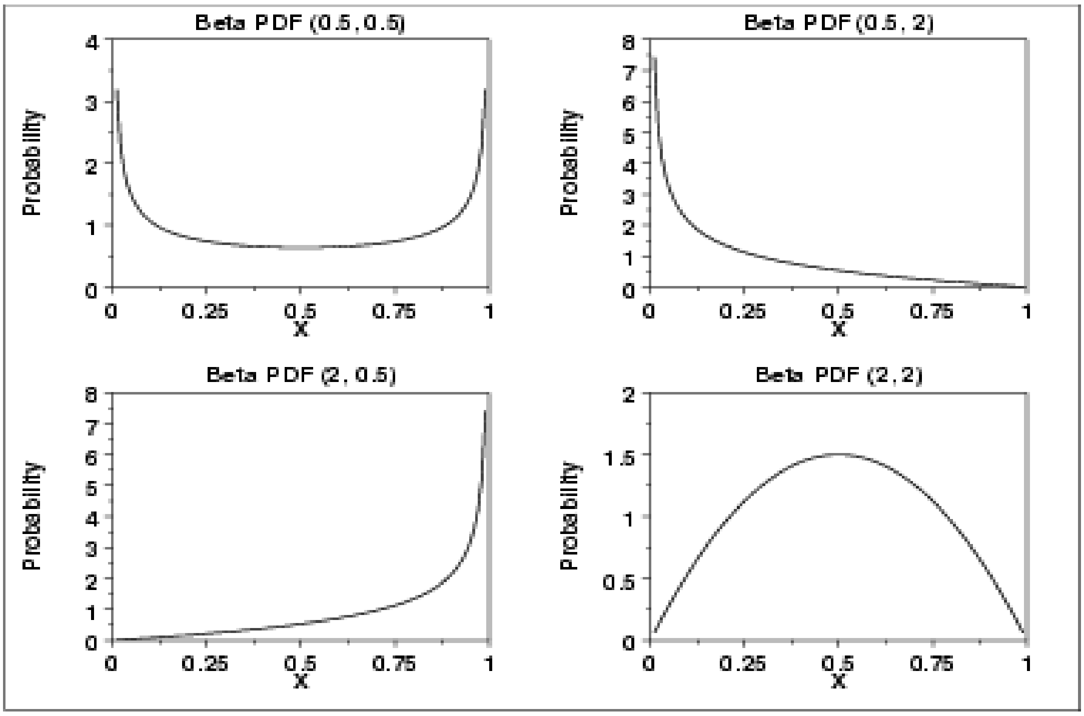


Fig. 2-4 the p.d.f of some Beta distributions

If X_1 and X_2 are independent random variables with a gamma distribution having parameters (α_1,θ) and (α_2,θ), then $X_1/(X_1+X_2)$ is a beta distribution variable with parameters (α_1,α_2). This can be derived as follows:

$$f(x_1,x_2)=\frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-(x_1+x_2)}x_1^{\alpha_1-1}x_2^{\alpha_2-1} \qquad (2.9)$$

And it is the gamma distribution

Let $u=x_1+x_2$ and $v=\frac{x_1}{x_1+x_2}$ then $\Rightarrow x_1=uv$ and $x_2=u(1-v)$

the Jacobian is: $J=\begin{vmatrix}\frac{\partial x_1}{\partial u} & \frac{\partial x_1}{\partial v}\\ \frac{\partial x_2}{\partial u} & \frac{\partial x_2}{\partial v}\end{vmatrix}=\begin{vmatrix}v & u\\ 1-v & -u\end{vmatrix}=-u$

so $g(u,v)dudv=f(x_1,x_2)dx_1dx_2=f(x_1,x_2)ududv$

$$\Rightarrow g(u,v)=\frac{u}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-(x_1+x_2)}x_1^{\alpha_1-1}x_2^{\alpha_2-1}=\frac{u}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-u}x_1^{\alpha_1-1}x_2^{\alpha_2-1}$$

$$=\frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-u}u^{\alpha_1+\alpha_2-1}v^{\alpha_1-1}(1-v)^{\alpha_2-1}$$

The ratio $X_1/(X_1+X_2)$ has the distribution

since gamma function $\Gamma(\alpha)=\int t^{\alpha}e^{-t}dt$

$$f(v)=f(\frac{x_1}{x_1+x_2})=\int_0^1 g(u,v)du=\int_0^1\frac{1}{\Gamma(\alpha_1)\Gamma(\alpha_2)}e^{-u}u^{\alpha_1+\alpha_2-1}v^{\alpha_1-1}(1-v)^{\alpha_2-1}du$$

$$=\frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)}v^{\alpha_1-1}(1-v)^{\alpha_2-1}$$

Which proves that $v=\frac{x_1}{x_1+x_2}$ is a beta distribution variable with parameters (α_1,α_2).

Introduction to Dirichlet distribution

In the model, the probabilities of the alleles at each locus are assumed to be Dirichlet distributed. The probability density of the Dirichlet distribution for variables $\vec{P} = (p_1, \cdots p_n)$ with parameters $\alpha = (\alpha_1, \cdots \alpha_n)$ is defined by:

$$p(\vec{P}) = Dirichlet(\vec{P}; \alpha) = \frac{1}{Z(\alpha)} \prod_{i=1}^{n} p_i^{\alpha_i - 1} \qquad (2.10)$$

In which $p_1, \cdots p_n > 0$; $\sum_{i=1}^{n} p_i = 1$ and $\alpha_1, \cdots \alpha_n > 0$.

The parameters α_i can be interpreted as prior observation counts for events governed by p_i. The normalization constant $Z(\alpha)$ is:

$$Z(\alpha) = \frac{\prod_{i=1}^{n} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{n} \alpha_i)} \qquad (2.11)$$

Let $\alpha_0 = \sum_{i=1}^{n} \alpha_i$, then the mean and variance of p_i is:

$$E(p_i) = \frac{\alpha_i}{\alpha_0} \text{ and } \operatorname{var}(p_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)} \qquad (2.12)$$

Proof: Since the marginal distribution of p_i is Beta distributed $(\alpha_i, \alpha_0 - \alpha_i)$:

$$E[p_i^r (1 - p_i)^s] = \frac{1}{B(\alpha_i, \alpha_0 - \alpha_i)} \int_0^1 p_i^{r + \alpha_1 - 1} (1 - p_i)^{s + \alpha_0 - \alpha_i - 1} dp_i = \frac{B(\alpha_i + r, \alpha_0 - \alpha_i + s)}{B(\alpha_i, \alpha_0 - \alpha_i)}$$

for $r = 1, s = 0$

$$E(p_i) = \frac{B(\alpha_i + 1, \alpha_0 - \alpha_i)}{B(\alpha_i, \alpha_0 - \alpha_i)} = \frac{\Gamma(\alpha_i + 1)}{\Gamma(\alpha_i)} \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_0 + 1)} = \frac{\alpha_i}{\alpha_0}$$

for $r = 2$, $s = 0$

$$E(p_i^2) = \frac{B(\alpha_i + 2, \alpha_0 - \alpha_i)}{B(\alpha_i, \alpha_0 - \alpha_i)} = \frac{(\alpha_i + 1)\alpha_i}{(\alpha_0 + 1)\alpha_0}$$

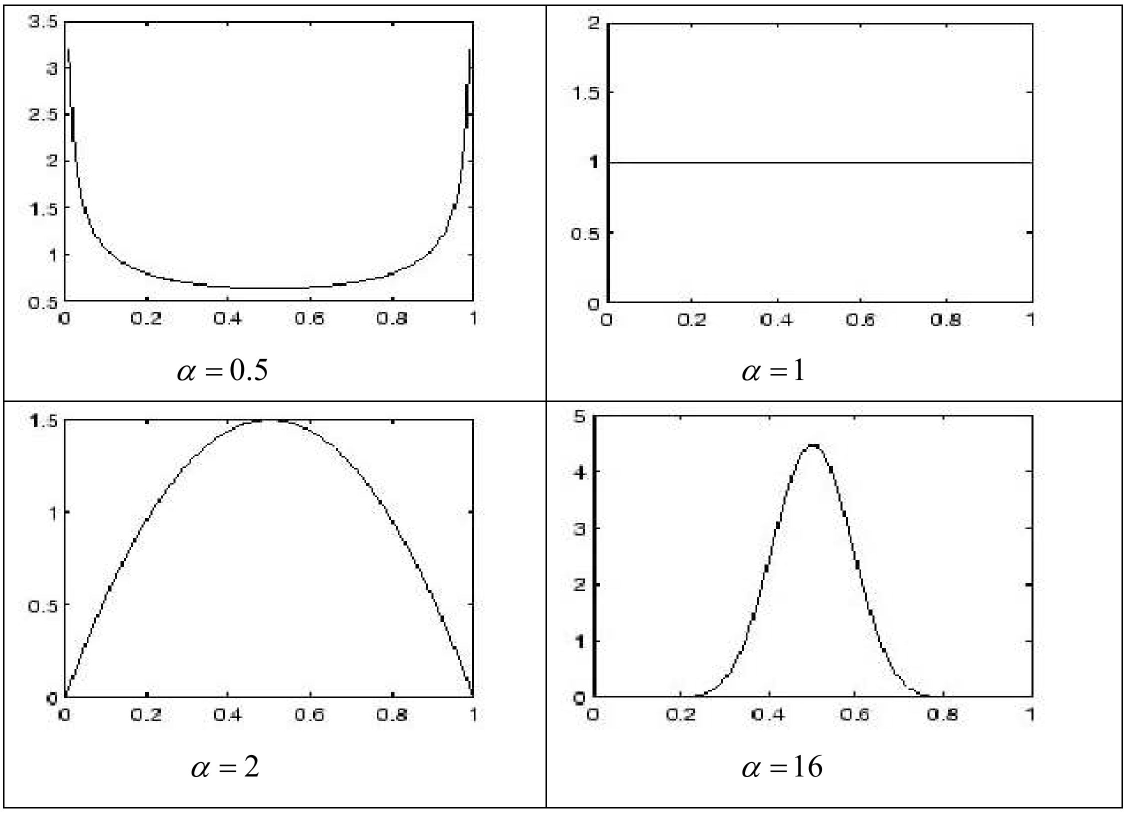
$$\mathrm{var}(p_i) = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\alpha_0 + 1)}$$

When $\alpha_i \to 0$, the distribution becomes no informative ($H(X) = -p(X)\log(P(X))$). The means of all the p_i stay the same if all α_i are scaled with the same multiplicative constant. The variances will, however, get smaller as the parameters α_i grows, $\alpha' = \tau \times \alpha$:

$$E(p_i') = \frac{\tau \times \alpha_i}{\sum_{i=1}^{n} \tau \times \alpha_i} = \frac{\tau \times \alpha_i}{\tau \times \sum_{i=1}^{n} \alpha_i} = \frac{\alpha_i}{\alpha_0} = E(p_i) \tag{2.13}$$

$$\mathrm{var}(p_i') = \frac{\tau^2 \alpha_i(\alpha_0 - \alpha_i)}{\tau^2 \alpha_0^2(\tau \times \alpha_0 + 1)} = \frac{\alpha_i(\alpha_0 - \alpha_i)}{\alpha_0^2(\tau \times \alpha_0 + 1)} \tag{2.14}$$

When τ goes up, $\mathrm{var}(p_i)$ will decreases. The pdfs of the Dirichlet distribution with certain parameter values are shown in the following figure [2].



$\alpha = 0.5$ $\alpha = 1$

$\alpha = 2$ $\alpha = 16$

Fig. 2-5 the p.d.f of some Dirichlet distributions ($\alpha_1 = \alpha_2 = \alpha$)

CHAPTER 3 Simulation of "Oslo bump"

The genetic method is powerful in species research and the lack of Hardy Weinberg equilibrium in the BCB population sampled at Barrow might indicate there being a mixture of two genetically distinct population compounds (Jorde et al, 2004). In this chapter, I will build gene pools by Dirichlet distribution and simulate the distribution of microsatellite alleles in bowhead whales.

The observation data and the distribution hypothesis

The 54 whale samples landed at Barrow in fall, with genetic measurement on 11 loci (Tv7, Tv11, Tv13, Tv17, Tv19, Tv20, Ga28, Ev1, Tv14, Ev104, Tv16) in the micro-satellite (Jorde et al, 2004), are studied. The whale samples are distributed in 10 years:

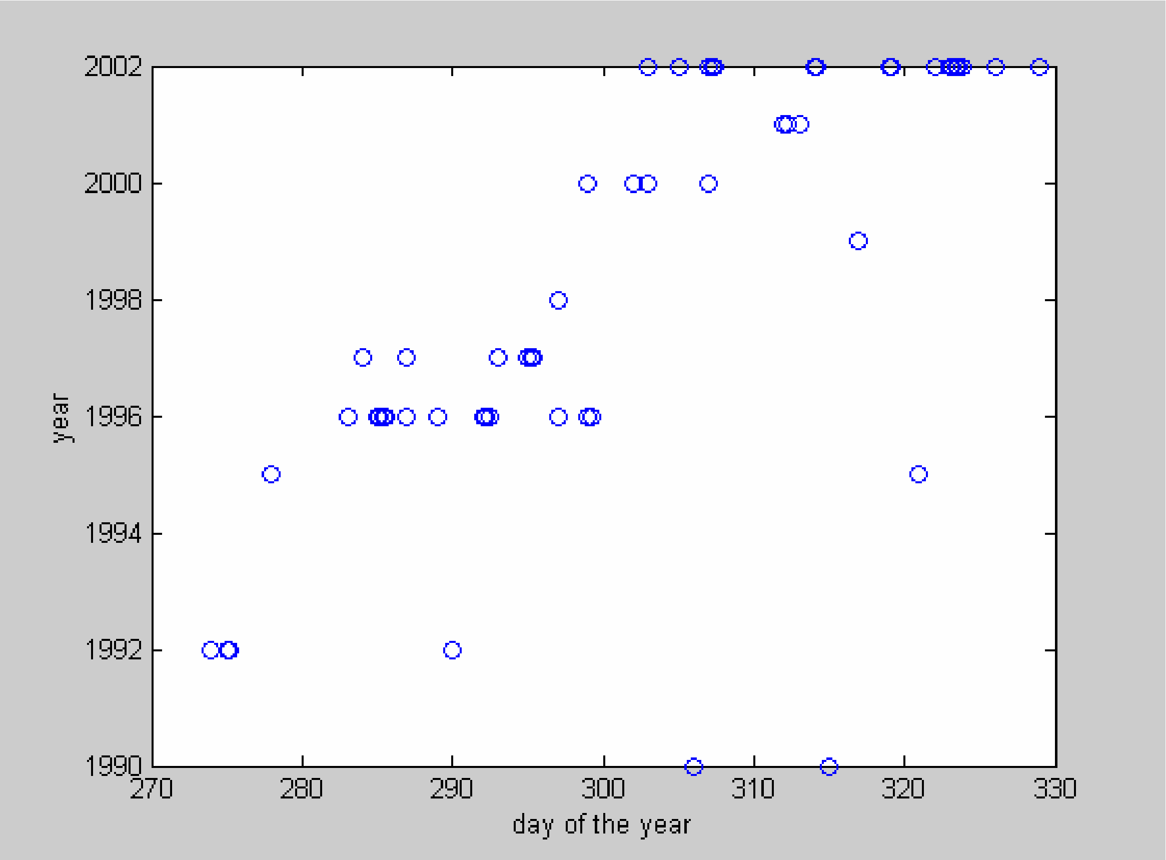


Fig. 3-1. Population allocation by year and day.

The whales' period of migration changes from year to year. To facilitate the comparison, the data are moved to a same middle day for every year:

$$Middle\_day(year) = \frac{(first\_day(year) + last\_day(year))}{2}$$

Fig. 3-2 is the figure after the data was moved to a same middle day. It shows that the migration period lasts about 15 to 30 days.

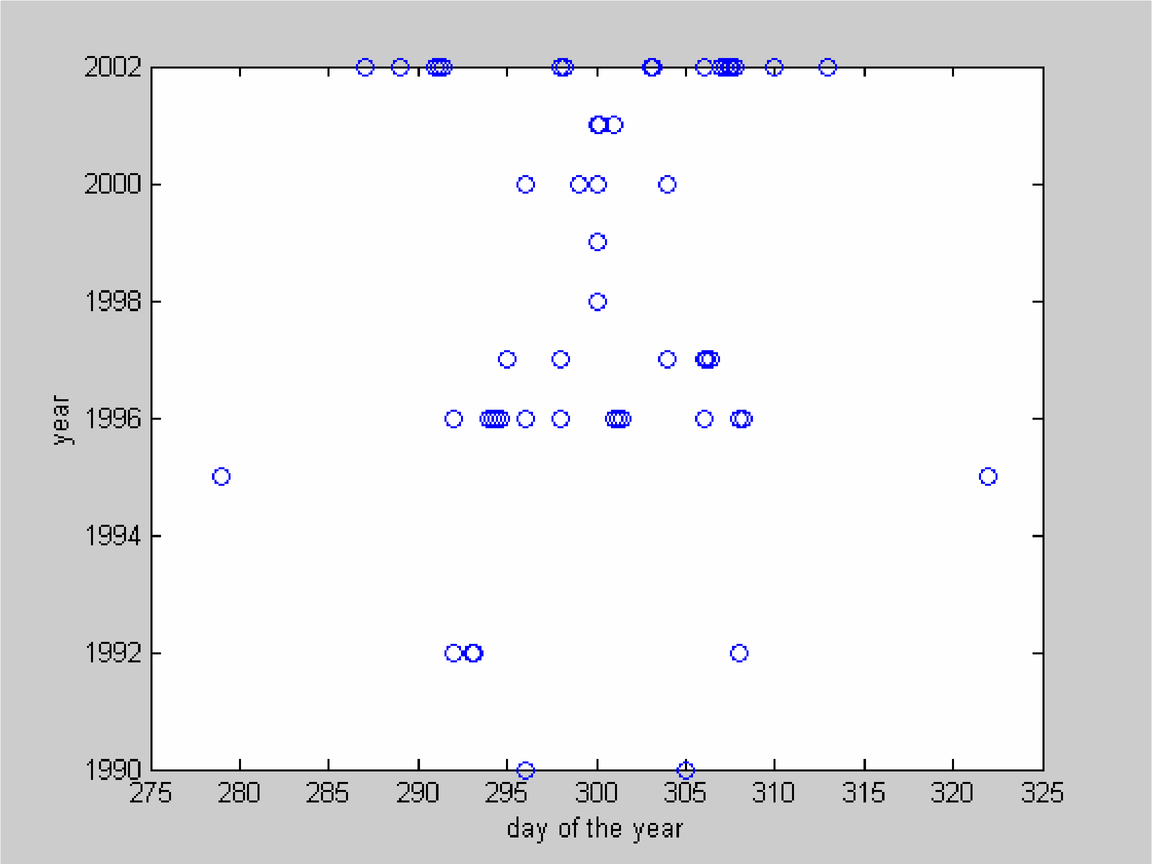


Fig. 3-2. Population allocation after moving to the same middle day.

It is difficult to find any pattern from the data in one year. To count the number of whales in each day, all the data in the same day was aggregated first and the result was depicted in Fig. 3-3.

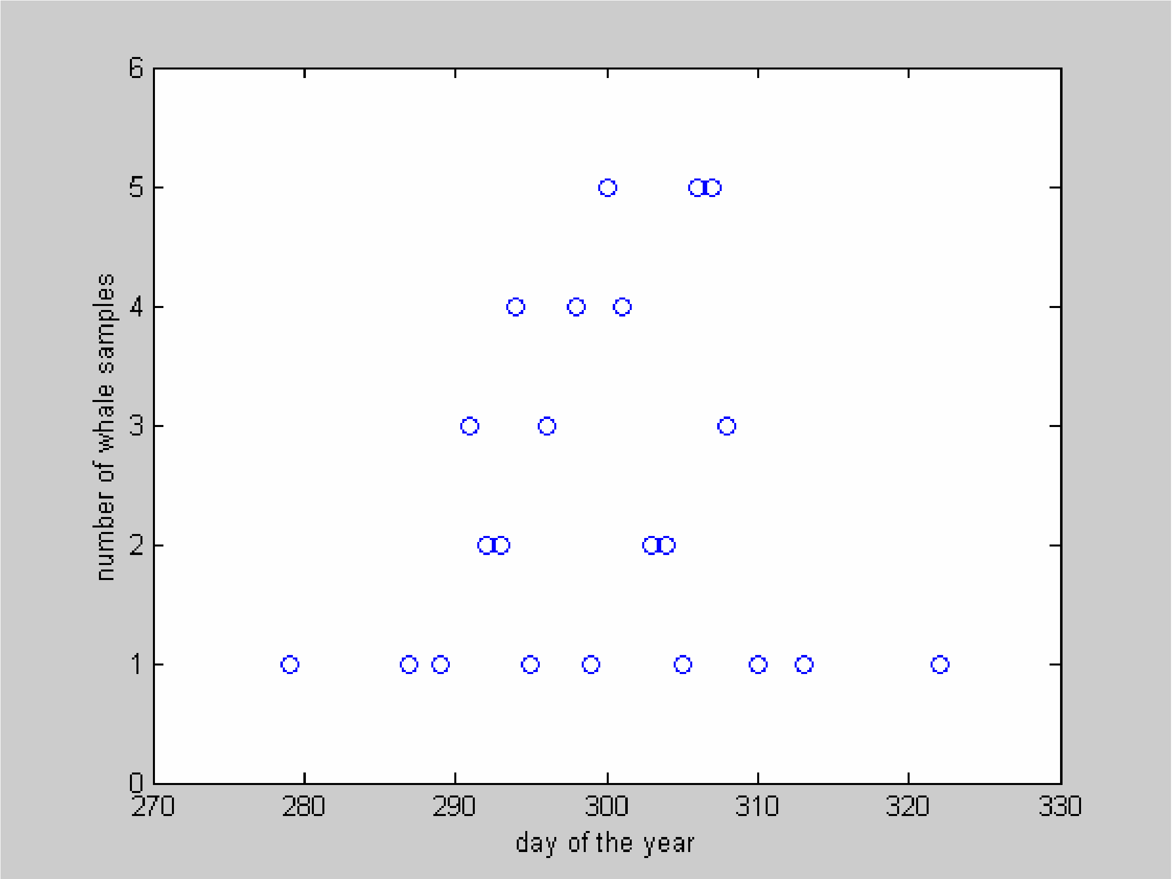


Fig. 3-3. The number of whales by day of year, centered within year

It shows that most of the whales come in a period from the 290th day to the 310th day in a year and there are two peak points ------ the 295th and the 305th day respectively, with a valley between them. However, at the 300th day a small peak also shows up. This is proved by the traditional Inuit knowledge, which says that both in spring and fall, bowhead whales pass Barrow in three pulses. If Chukchi circuit hypothesis is right, we can imagine that there are two E groups pass Barrow strait around the 295th and the 305th day of the year respectively, while a small W group pass Barrow strait between them.

Assuming the pulses having a normal distribution over days, in the sense that the expected number of whales passing Barrow in a particular day is proportional to:

$$D = \frac{n}{\sqrt{2\pi} \times \sigma} \exp[-\frac{(day - \mu)^2}{2 \times \sigma^2}]$$

The mean μ should be the peak day, the variance σ indicates how long the whale pulse lasts and the n stands for how large the pulse is. A fitting hypothesis is given first to the observing data in Fig. 3-3:

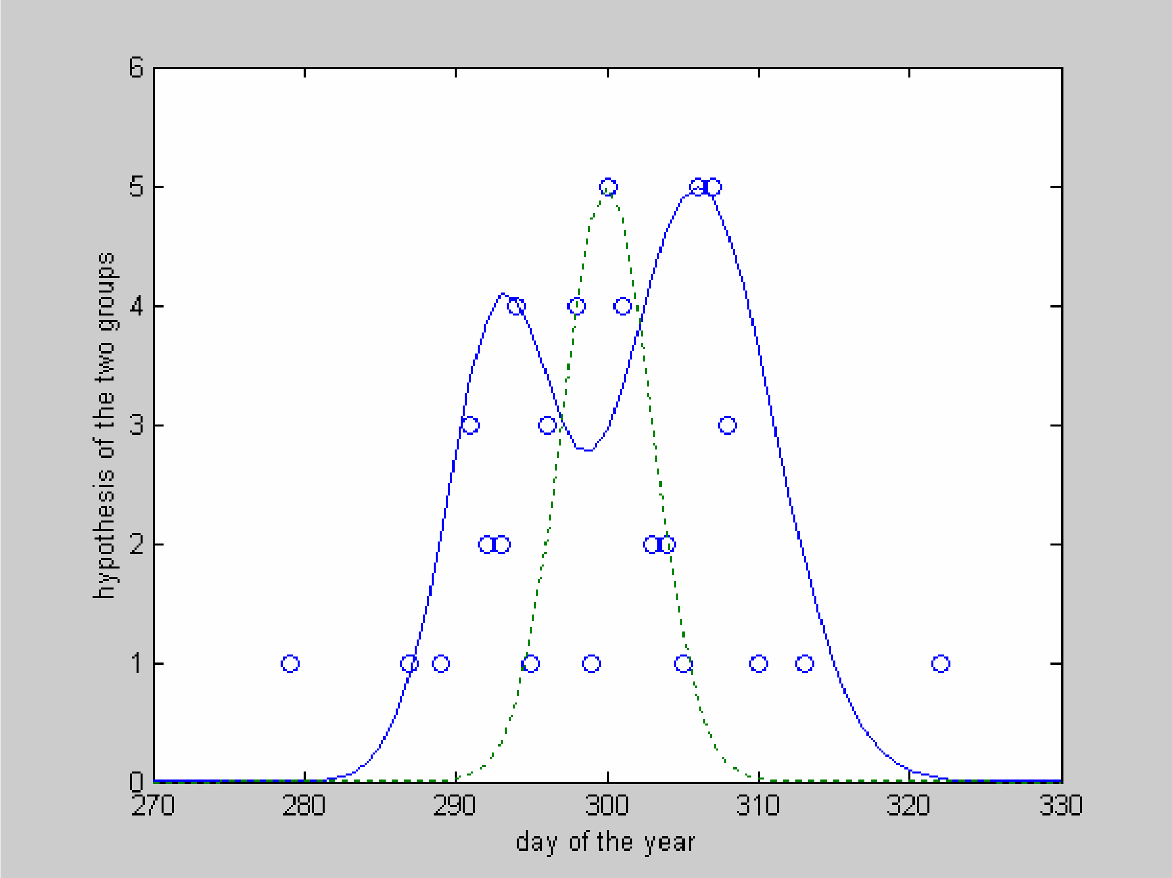


Fig. 3-4. Sample distribution and hypothetical composition

E-whales come in to pulses, solid line

W-whales come in a middle pulse, broken line

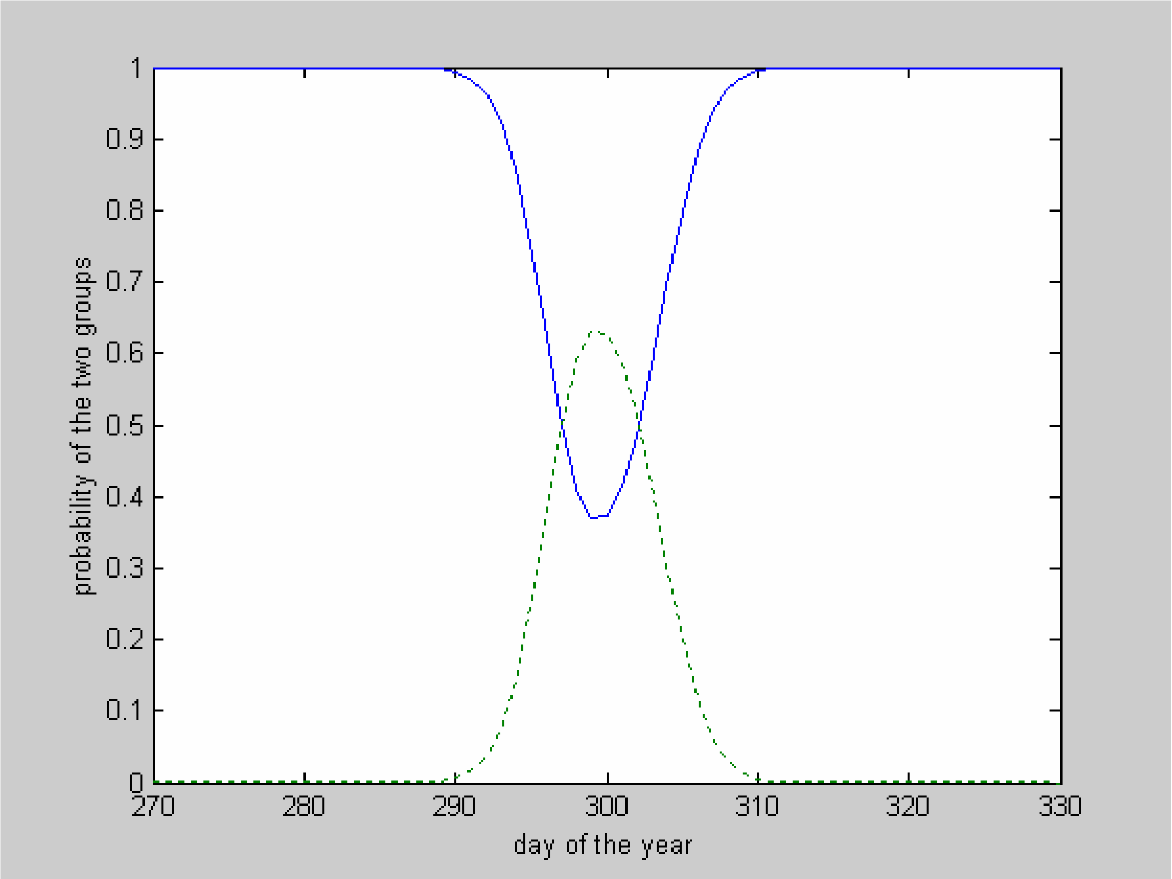


Fig. 3-5. The probability $P_E(t)$ (the solid line) and $P_W(t)$ (the broken line)

Based on the distribution of E and W-groups, we can judge which group that every whale sample belongs to. The probability of a whale passing Barrow at day t belongs to E-group can be calculated by the following function:

$$p_E(t) = \frac{D_E(t)}{D_E(t) + D_W(t)} \quad ,$$

and the corresponding probability of a whale belongs to W-group is:

$$p_W(t) = 1 - p_E(t)\,.$$

Then for each sample at day t, a random number which is uniformly distributed in the interval (0,1) is given by the computer and compared to $P_W(t)$. When the random number is larger than $P_W(t)$, the simulated whale is assigned to be in E-group, otherwise the whale is assigned to be in W-group. Fig. 3-6 depicts the belonging of each whale sample by using the above category method.

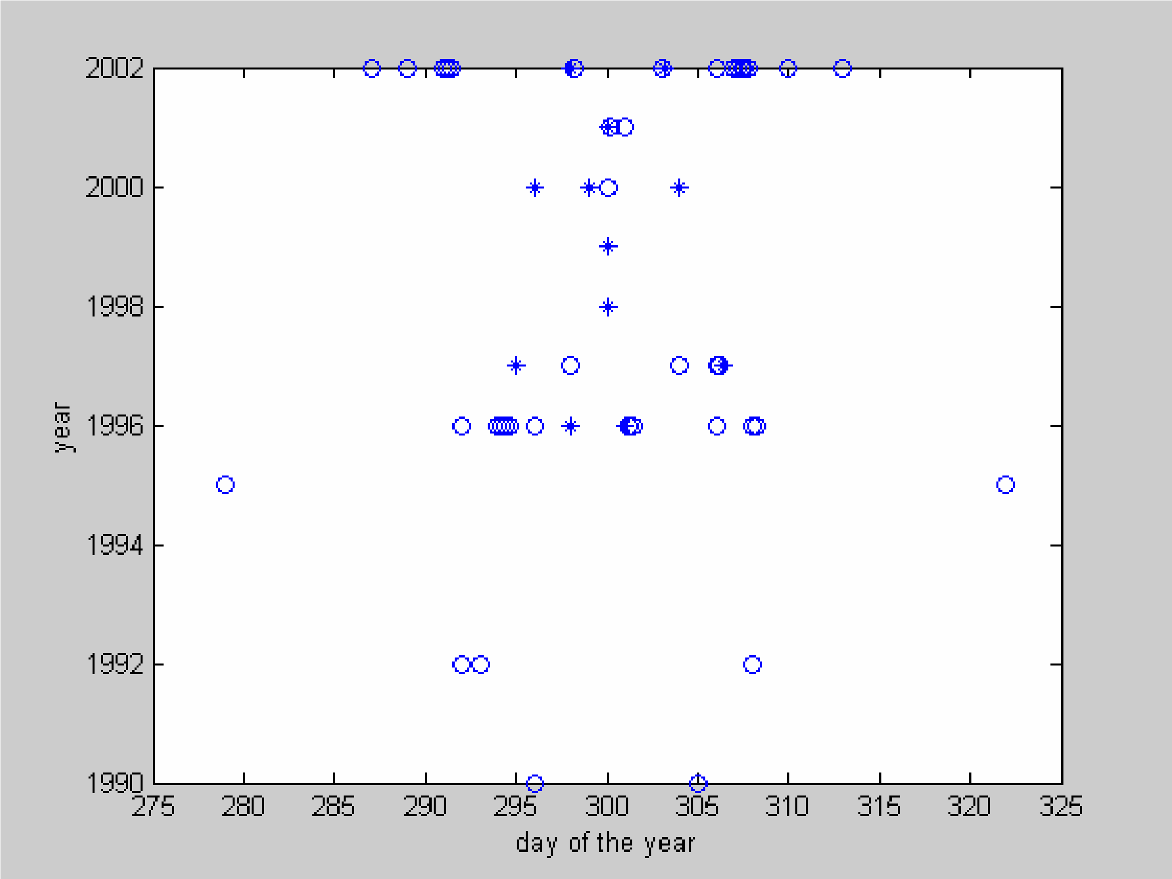


Fig. 3-6. E and W population

(E-group is described by “o” and W-group is described by “*”)

From this figure, we can see the W-group mainly distributed from 295 to 305, while some of the E-group is also distributed in this interval.

The program is “**Whale_DataPro_Main.m**”

The construction of gene pools

The model used in this paper should be stochastic with respect to the genetic composition in the two hypothetical populations E and W, and thus also with respect to the genetics of the sampled individuals in each simulation run. In each run, the gene pools in E and W are respectively generated by two draws from a common Dirichlet distribution. Each of the sampled whales is then randomly assigned with genes drawn from the E-pool with probability $p_E(t)$ and from W-pool with probability $p_W(t)$, in which t is the day of sampling.

At each locus of the individual's gene, the distribution of the alleles is supposed to follow a Dirichlet distribution across the populations. The parameters of the Dirichlet distribution can be interpreted as "prior observation counts" and are estimated from the data as the observed frequencies coupled with a speculative value of α_i:

$$\alpha_i = \frac{frequency_i}{\sum_i frequency_i}, \quad i = 1,2,\cdots m$$

m is the number of unique alleles in the locus. The following table is an example for one of the locus, say TV16:

allele	184	186	190	192
frequency	7	160	57	10
α_i	0.0299	0.6838	0.2436	0.0427

Table. 3-1. Probability α_i at locus TV16

The allele distribution in a gene pool is for each locus drawn from a Dirichlet distribution:

$$p(P) = Dirichlet(P;\alpha) = \frac{1}{Z(\alpha)} \prod_{i=1}^{n} p_i^{\alpha_i - 1}$$

where $p_1, \cdots p_n > 0; \sum_{i=1}^{n} p_i = 1, \ \alpha_1, \cdots \alpha_n > 0$.

and $Z(\alpha) = \frac{\prod_{i=1}^{n} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{n} \alpha_i)}$

Thus for each locus a Dirichlet distribution is constructed. Two independent Dirichlet gene pools are constructed for E- and W -group respectively. The gene pool for E group is generated as one draw from this distribution, as is the gene pool for W group.

The method to assign alleles to individual whales is essentially the same as each individual is assigned E- or W-group in the simulation. Since the whale is diploid, it inherits one gene from its mother and another gene from its father, which are supposed to belong to the same pool. After population assignment is done, two independent draws (one for each parent) from the same gene pool is carried out for each locus.

The genetic difference between the two individuals

The observed pattern called the Oslo bump was recognized when pair wise genetic distance was plotted against days between catch within year. For two sampled whales taken d days apart, a measure of genetic distance a aggregated over loci was calculated. The plot of a versus d showed a bump centred at d is about 7 days. Now, I will explain the genetic distance a.

Consider a pair of individuals i and j. Since the whale is diploid, then at each locus, comparing each of the two genes in individual i to those in individual j makes up four comparisons. There can be 0, 1, 2 or 4 matches, which are summarized in Table 3-2.

$I_{ij} = 4$	Whale i (a,a) Whale j (a,a)	$I_{ij} = 1$	Whale i (a,b) Whale j (a,c)
$I_{ij} = 2$	Whale i (a,b) Whale j (a,a)	$I_{ij} = 0$	Whale i (a,b) Whale j (c,d)
$I_{ij} = 2$	Whale i (a,b) Whale j (a,b)		

Table. 3-2. Genetic difference

The number of matches, I_{ijk}, for whale pair (i,j) at locus k is contrasted to the locus-specific frequency of identity among genes within individuals. By averaging over loci with non-missing values, a measure of pair-wise genetic difference between the two individuals can be obtained (Rousset, 2000):

$$a_{ij} = \frac{1}{L_{ij}} \sum_k \frac{h_k - I_{ijk}/4}{1 - h_k}$$

in which L_{ij} is the number of loci that were successfully scored in both individuals i and j. h_k is the average homozygosity, which is the average of all individuals in the questioned sample at a single locus:

$$h_k = \sum_{i=1}^{n} q_i^2$$

where q_i is the empirical frequency of the allele in the samples. For example, if 3 of the total 54 simulated whales are of genotype (6 6), (6 7), (6 6) at locus Tv7, while none of the remaining 51 have allele 6, then the empirical frequency of the 6th allele of Tv7 is 5/108. We can prove that the expectation of $I_{ij}/4$ equals h, that is, $E(I_{ij}/4) = h$ by **Hardy-Weinberg** law under random mating in a mixed population. The probability of one of the four combinations to be "aa" is:

$$p_a^2 + 2p_a(1 - p_a) + \sum_{b \neq a} p_b^2 + 2 \sum_{b \neq c \neq a} p_b p_c \quad = \sum_i p_i^2 \,, \quad \text{where } \sum_{i=1}^{n} p_i = 1 .$$

Since there are four possible combinations in each situation, we can get:

$$E(I_{ij}) = 4\sum_i p_i^2 = 4h$$

Sometimes, h_k may equal 1 when α_i is small. This should not be a surprise since the probability is quite high that one allele will be predominant, which means almost all the members in the species will have this allele. To avoid computational error in the program, a judgment was introduced to let a = 0 when $h_k = 1$.

Simulation Results and Analysis

The whole program can be expressed by the following structure diagram:

For different distribution hypothesis p_E and p_W:

For every τ, do the following 100 times:

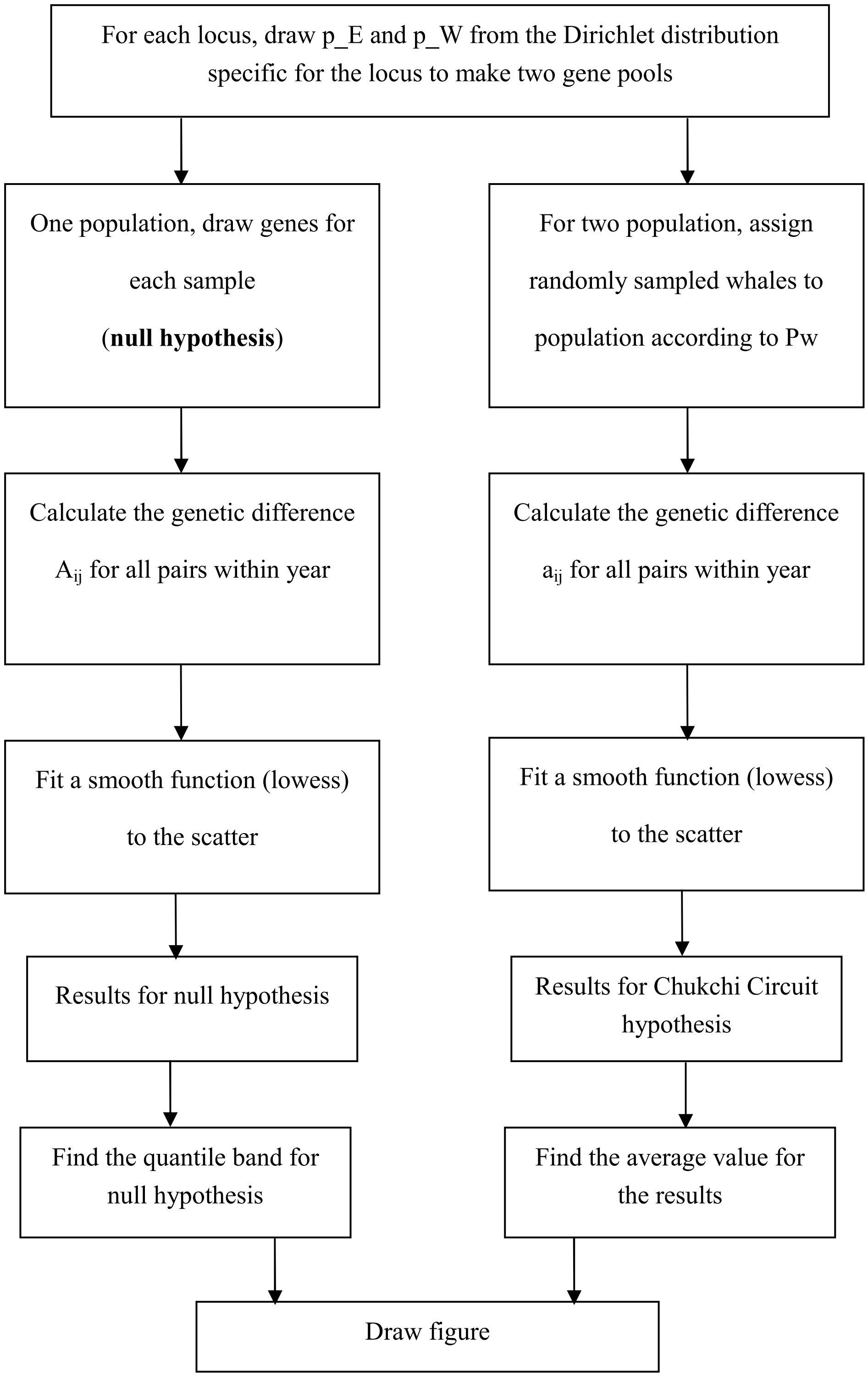


Fig. 3-7. Program structure diagram.

As we have mentioned before, when α' was defined as $\alpha' = \tau \times \alpha$, $\sum \alpha_{ik} = 1$ for each locus k. Different values of τ leads to different results. In the present case, τ was first set to be 0.1, 0.5, 1.0, 1.5 and 2.0. The results in the format of the scatter of pair wise genetic difference by days apart are showed in following figures:

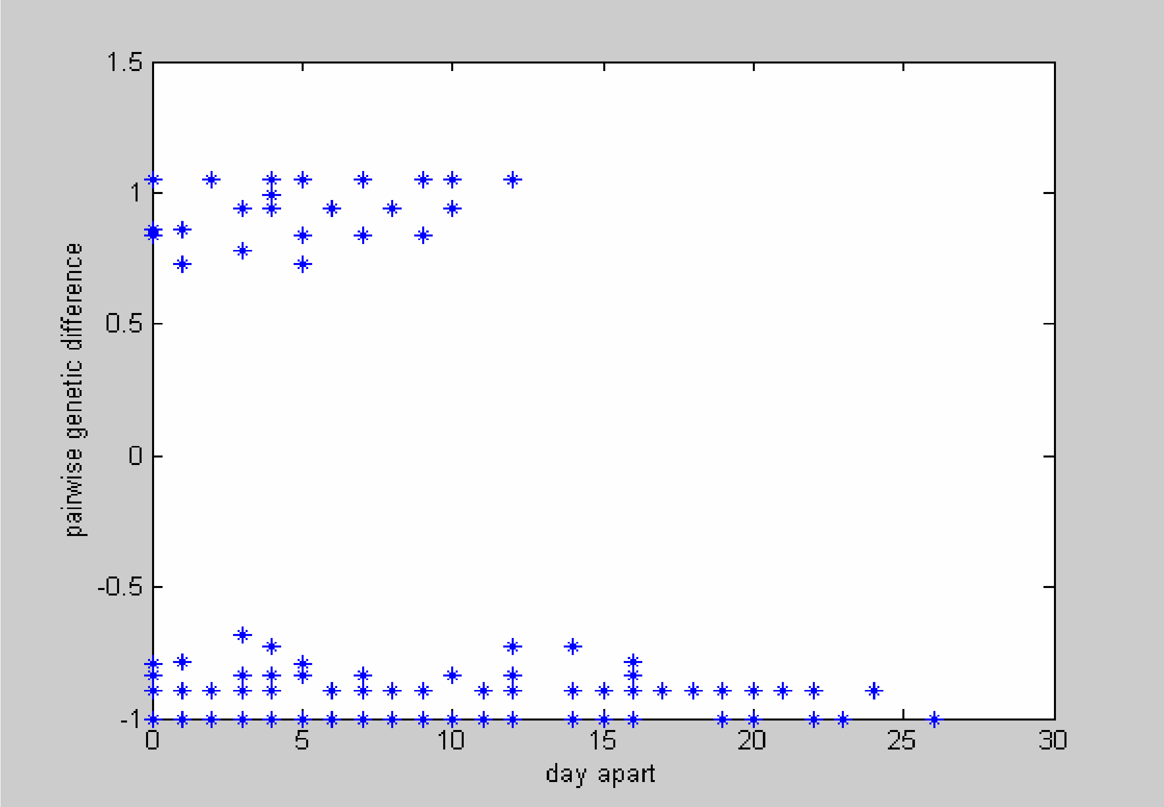


Fig. 3-8. Simulation result when τ =0.1

Two populations, days apart.

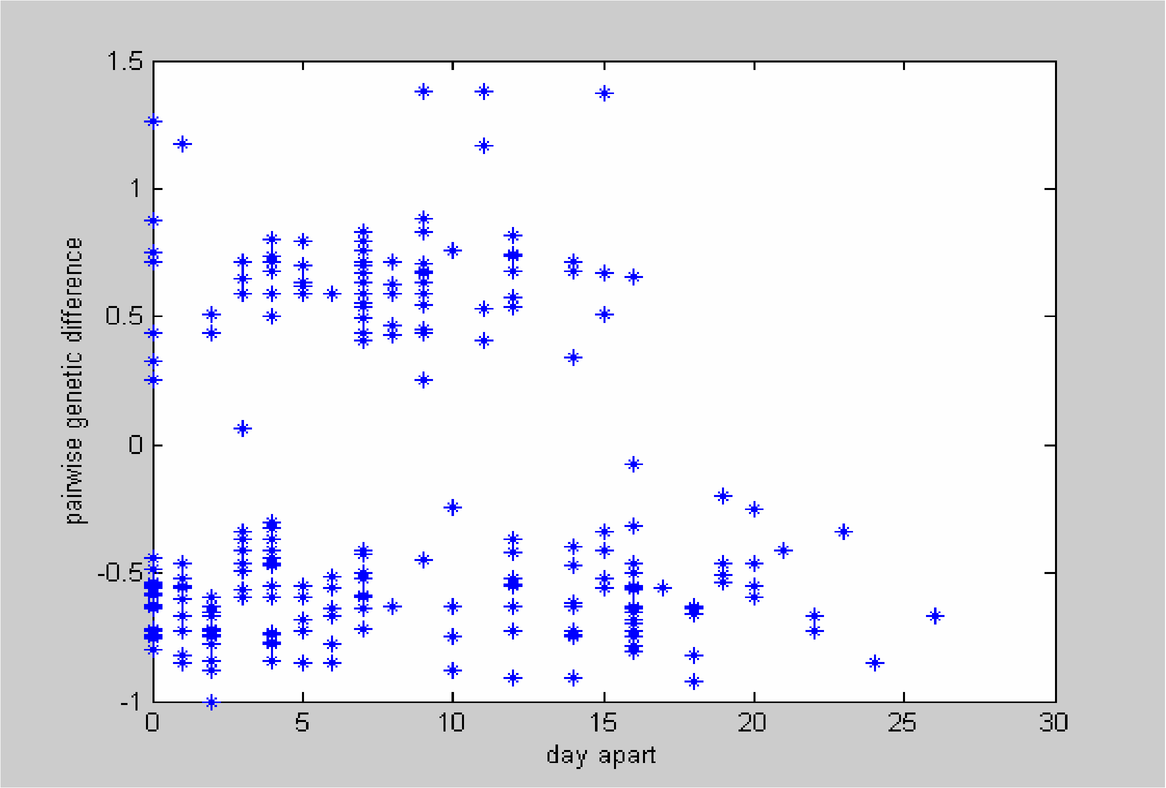


Fig. 3-9. Simulation result when τ =0.5

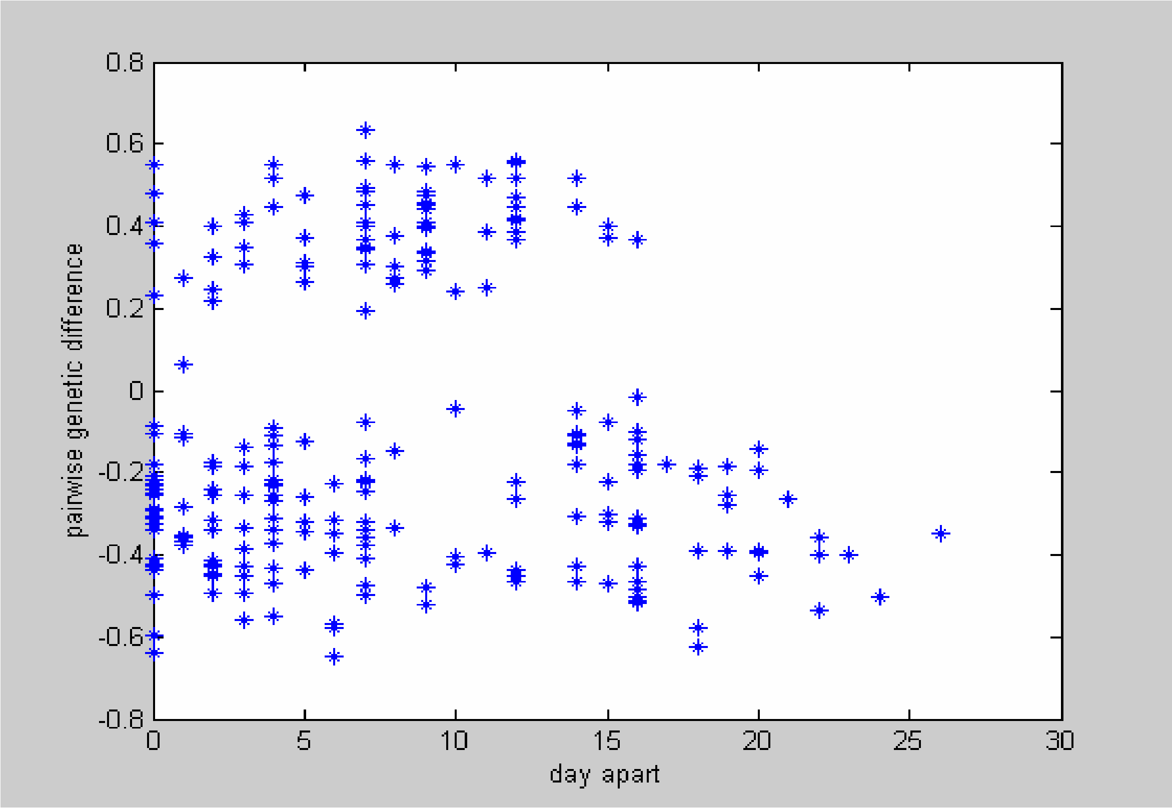


Fig. 3-10. Simulation result when τ =1.0

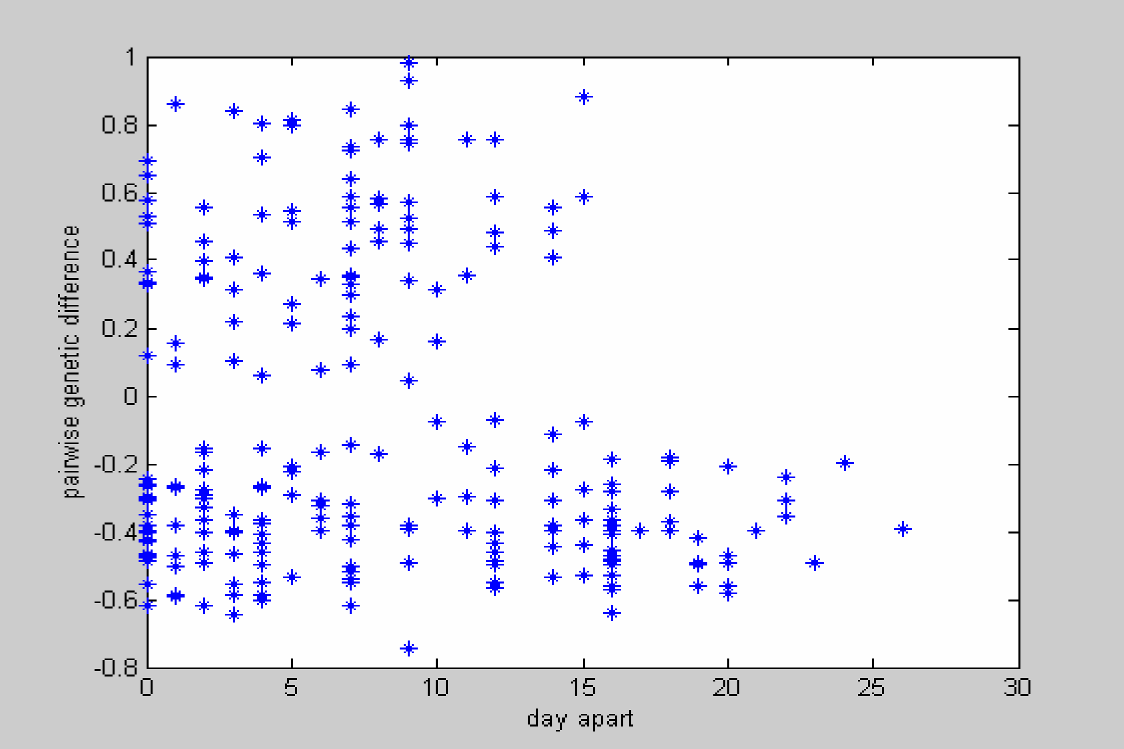


Fig. 3-11. Simulation result when τ =2.0

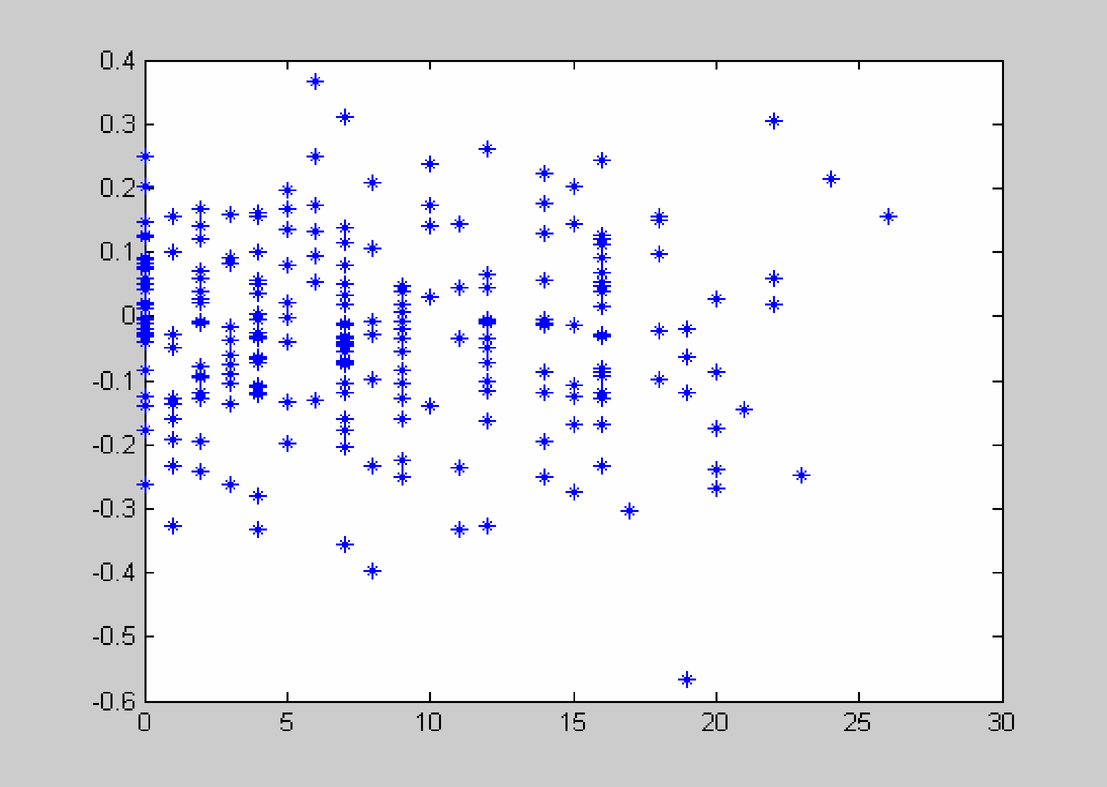


Fig. 3-12. Simulation result when τ =10

The program is **"Whale_GenePro_Main.m"**

From the figures above, it is clear that there are more positive points in the range of 3 to 12 days. When the interval exceeds 20 days or diminishes to 0, the values of the pair-wise difference becomes remarkable small. When day apart is 0, the average of A_{ij} is negative. According to the definition of A_{ij}:

$$A_{ij} = \frac{1}{L_{ij}} \sum_k \frac{h_k - I_{ijk}/4}{1 - h_k}$$

where $h_k = \sum q_i^2$, and $q_i \geq 0$, $\sum q_i = 1$, we can prove $h_k < 1$. So Aij will be small when I_{ij} is large. The results indicate I_{ij} is large when days apart are 0 or 20 days and we can conclude that there are more genetic matches then. When we calculate h, individuals from two groups were considered because of the hypothesis. This causes h to be smaller than the expectation of I_{ij} when the day apart is 0. The average of A_{ij} is positive when day apart is in the range of 3 to 12 days and this indicates that I_{ij} is small there are more differences between pair wise genes.

From the variance formula (2.14), the variability in the Dirichlet distribution is large for small values of τ. In the case the two draws from the Dirichlet distribution tend to be much different, an E-whale will then tend to be genetically much different from an W-whale. Pairs of whales are most likely of type EW when days apart is about one week. This explains Fig. 3-8.

When τ becomes larger, the values of the pair wise genetic difference becomes smaller as shown in the figures above. The changing of the genetic difference with τ can be seen clearly from the following Fig. 3-13, in which the standard deviation of a_{ij} was plotted against τ.

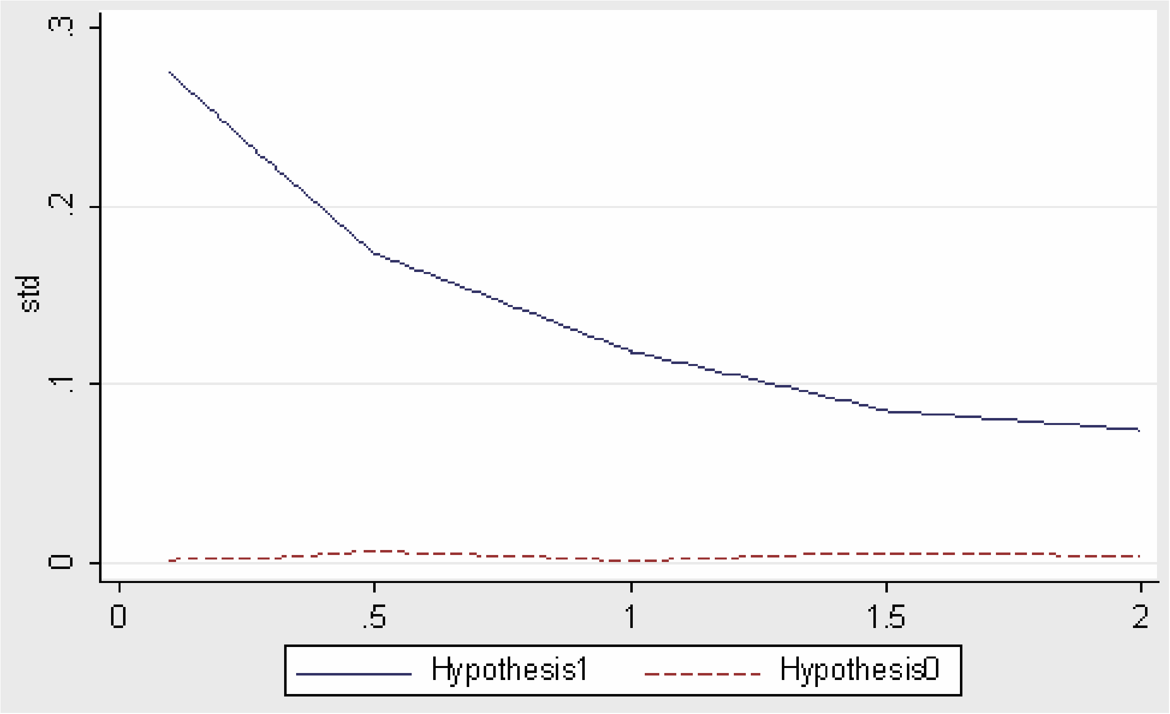


Fig. 3-13. The standard deviation v.s different τ
(hypothesis 0 means null hypothesis)

To avoid accident result, the pairwise genetic difference was simulated for 100 times at each τ and the average value was obtained. To demonstrate the general trend of the data, lowess function is used to get a smooth line. Fig. 3-14 depicts the results of pair wise genetic difference averaged for 100 times' simulation when τ =1.0.

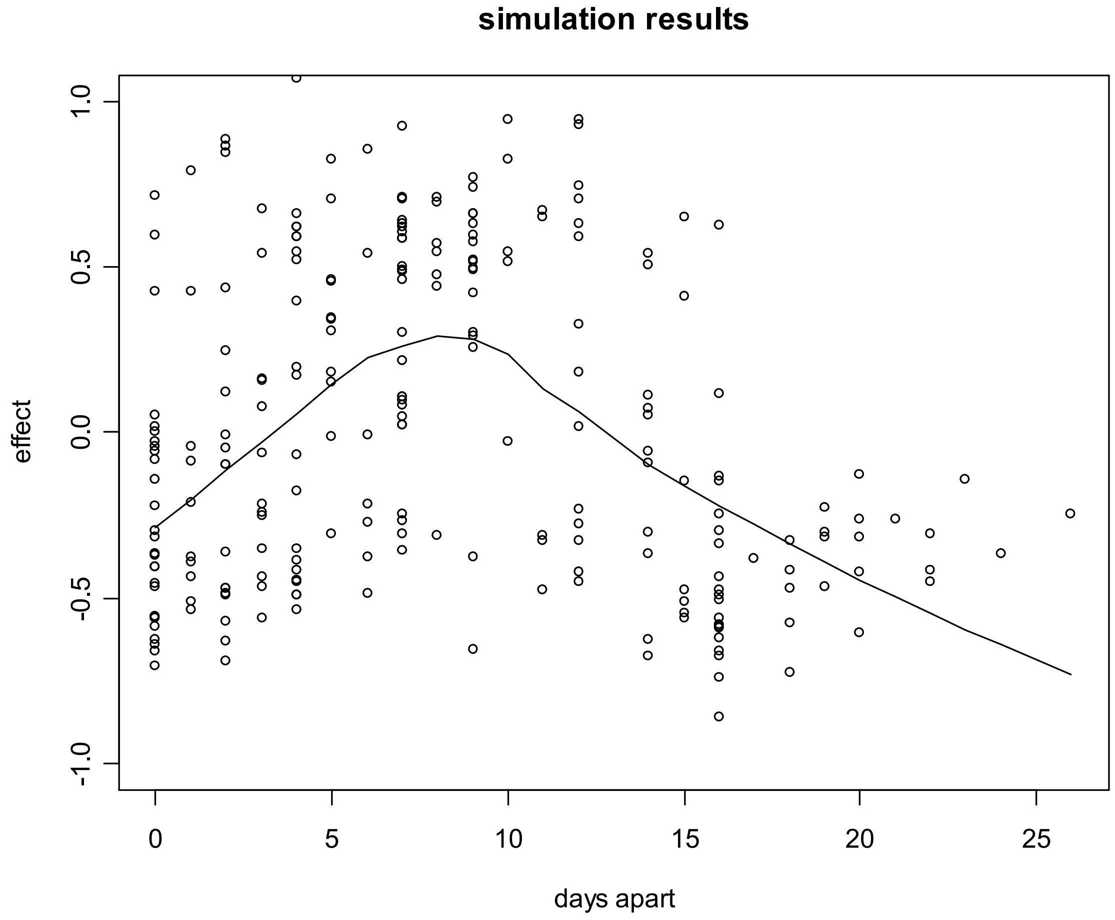


Fig. 3-14. Average result of 100 times simulation when τ =1.0.

(The solid line shows the fitness result by using the lowess function)

To compare the results from the null hypothesis and Chukchi circuit hypothesis, each τ was simulated 100 times under both hypothesises. The trend from Chukchi circuit hypothesis was compared with the 50%, 90%, 95% and 99% quantile of result from null hypothesis at each τ, that is the pointwise simulation. The following figures are the pointwise simulation results for τ =0.1, 0.3, 0.5, 1.0, 1.5 and 2.0 respectively:

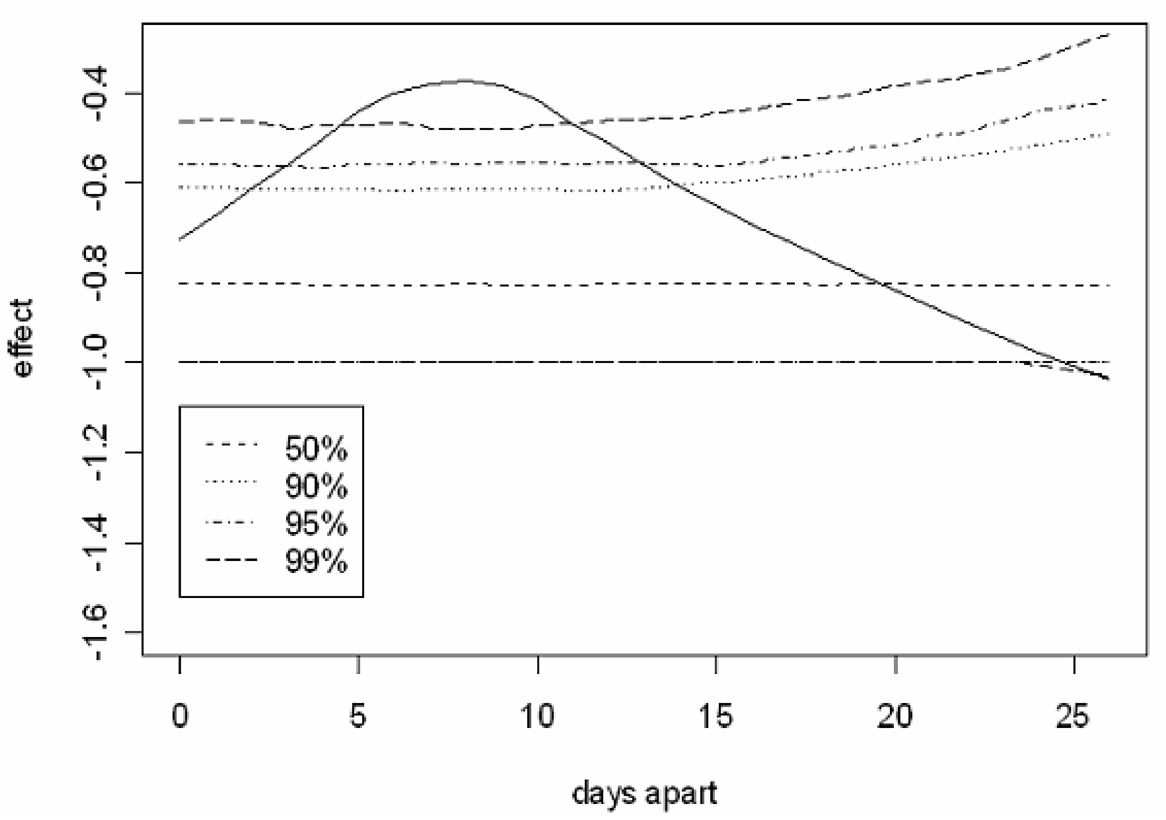


Fig. 3-15 Simulation result with pointwise band when τ =0.1

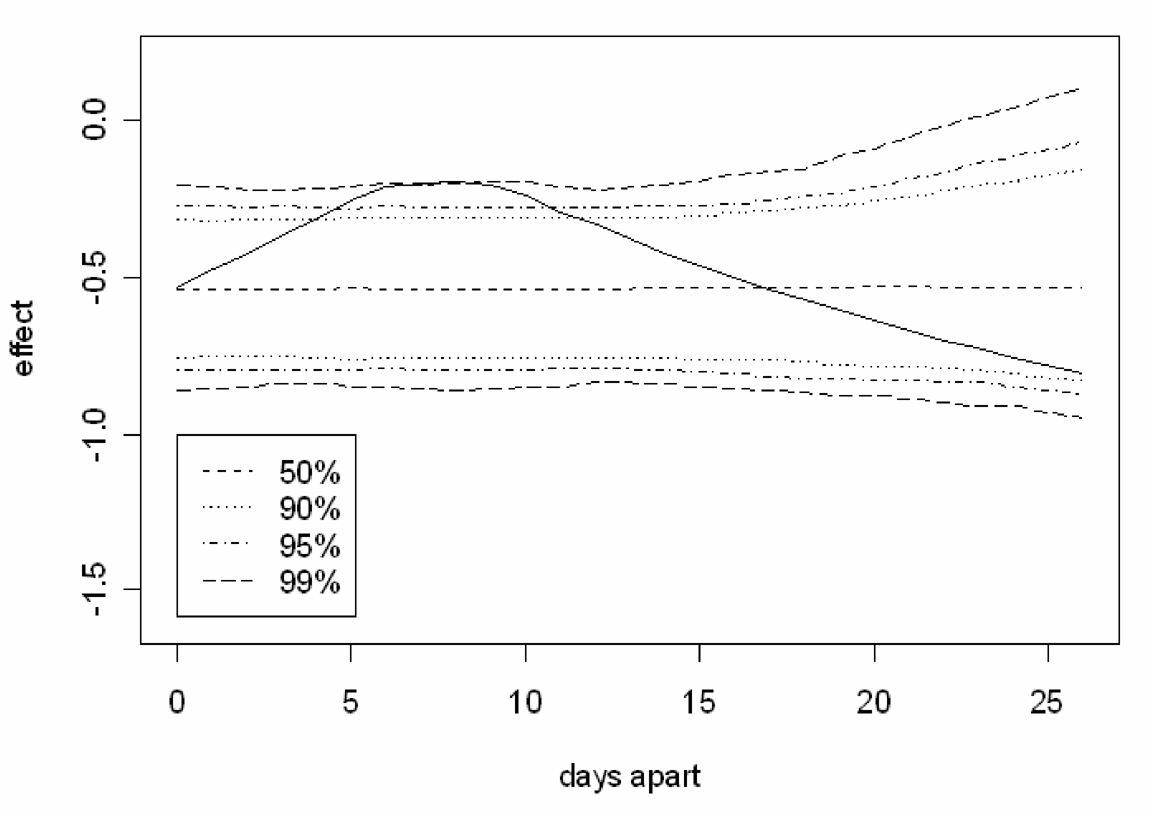


Fig. 3-16 Simulation result with pointwise band when τ =0.3

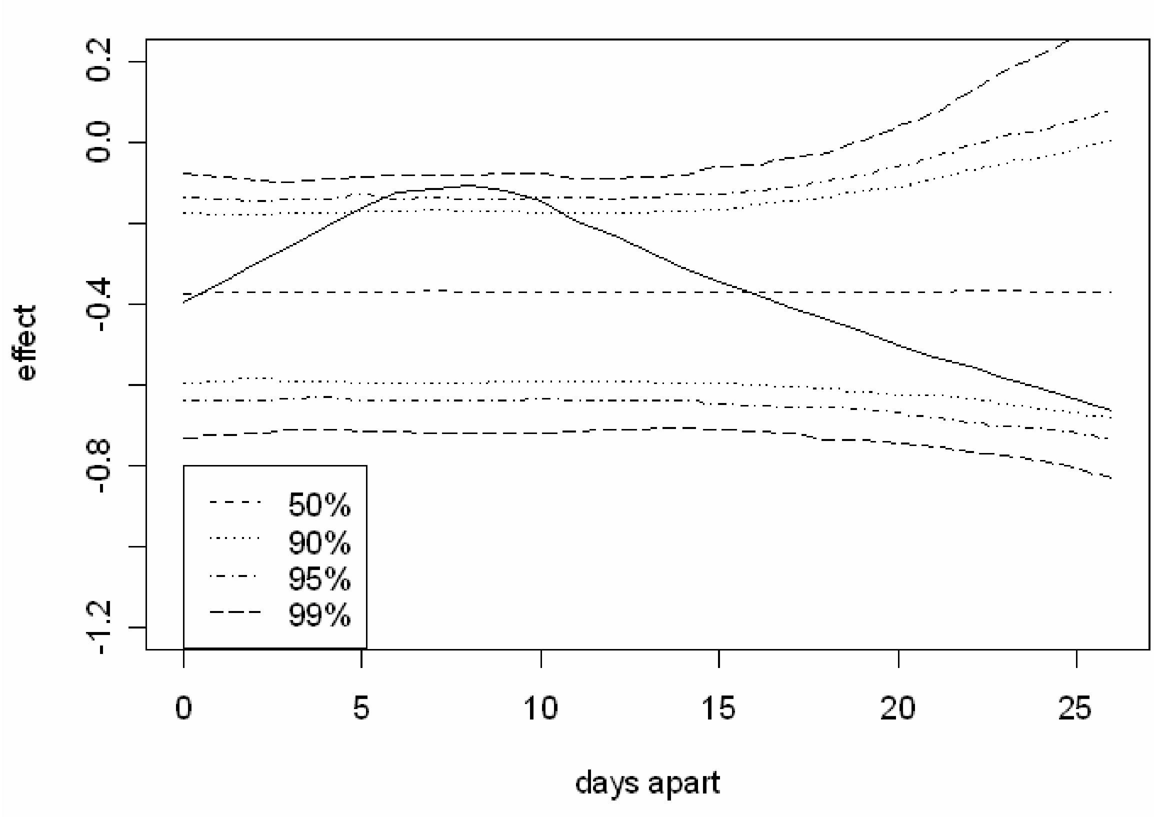


Fig. 3-17 Simulation result with pointwise band when τ =0.5

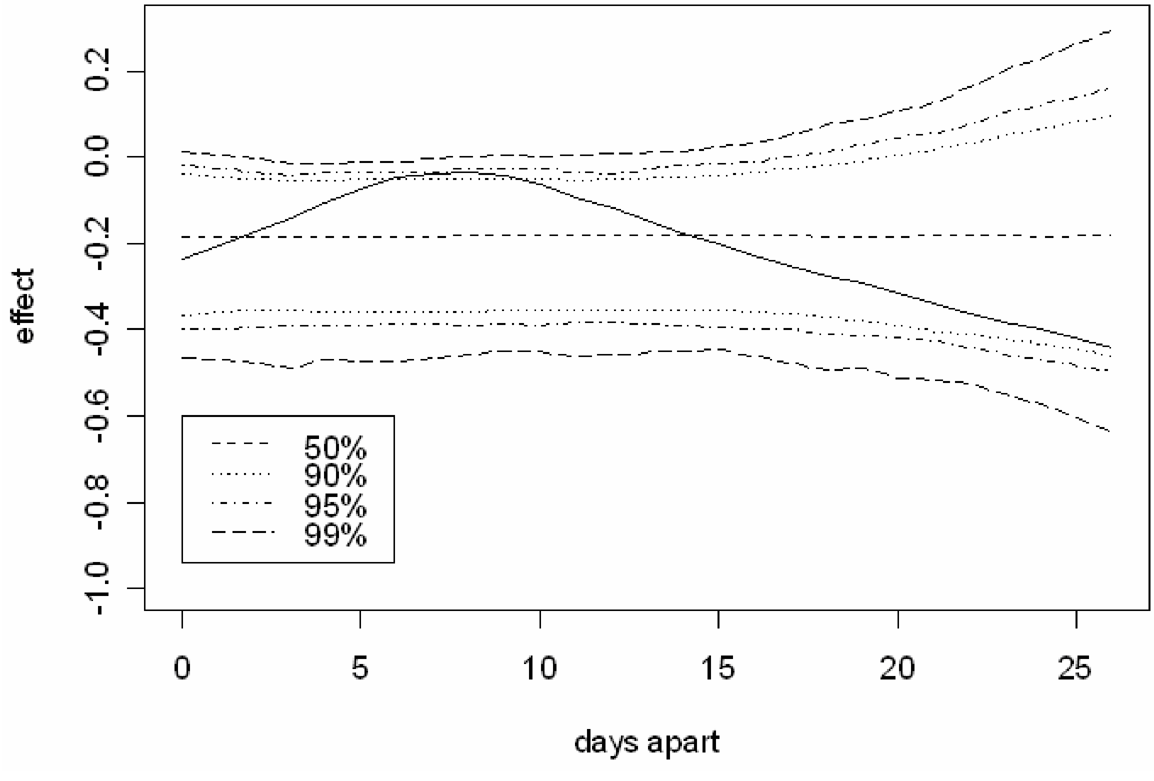


Fig. 3-18 Simulation result with pointwise band when τ =1.0

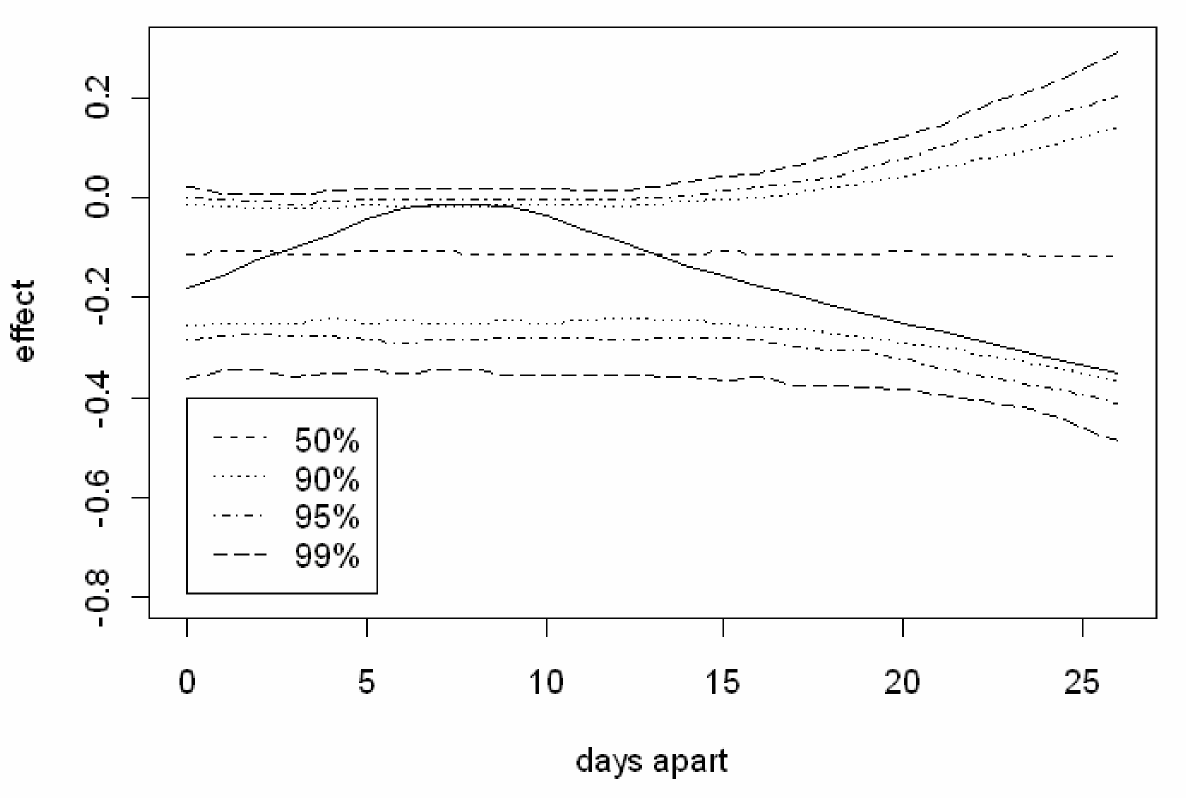


Fig. 3-19 Simulation result with pointwise band when τ =1.5

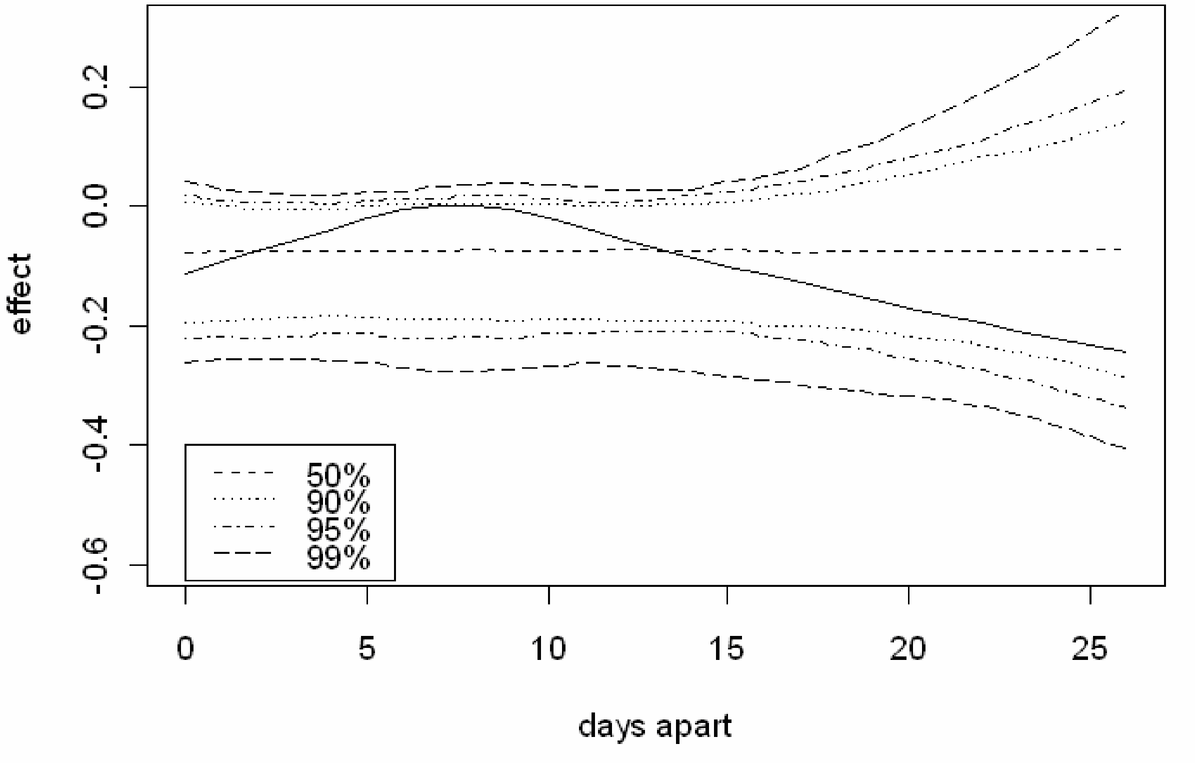


Fig. 3-20 Simulation result with pointwise band when τ =2.0

The program is **"test_1.R"**

From the figures; first we can see that the variance of the Chukchi circuit hypothesis is much bigger than that of null hypothesis. The genetic differences are obviously larger when paired samples were 5-11 days apart than they were 0-5 or 12-25 days apart.

Second, the curve of Chukchi circuit hypothesis becomes flat when τ becomes big.

Third, the null hypothesis can not be rejected at 95% significance level when $\tau > 1.0$. So it is important to study how much the genetic variance should be. In genetics, dissimilarity indexes (Fst) is used to indicate the molecular variance. It is calculated in the following way:

```
For locus $i = 1,\cdots,N$ {
    For allele $j = 1,\cdots n_i$ in locus i {
        $x_{ijk}$ = frequency of allele j in locus I for group k=E,W
        $Fst_{ij} = \sum_k (x_{ijk} - \bar{x}_{ij})^2 / \bar{x}_{ij}(1-\bar{x}_{ij})$, where $\bar{x}_{ij} = \frac{1}{2}(x_{ijE} + x_{ijW})$
        $Fst_i = \frac{1}{n_i}\sum_{j=1}^{n_i} Fst_{ij}$
        $Fst = \frac{1}{N}\sum_{i=1}^{N} Fst_i$
}}
```

There should be a functional relationship $Fst = g(\tau)$. A program was built up and the simulation result is shown in the following Fig. 3-21. In the paper (LeDuc et al. 2005), they estimated the genetic difference between these two populations and give Fst=0.062.

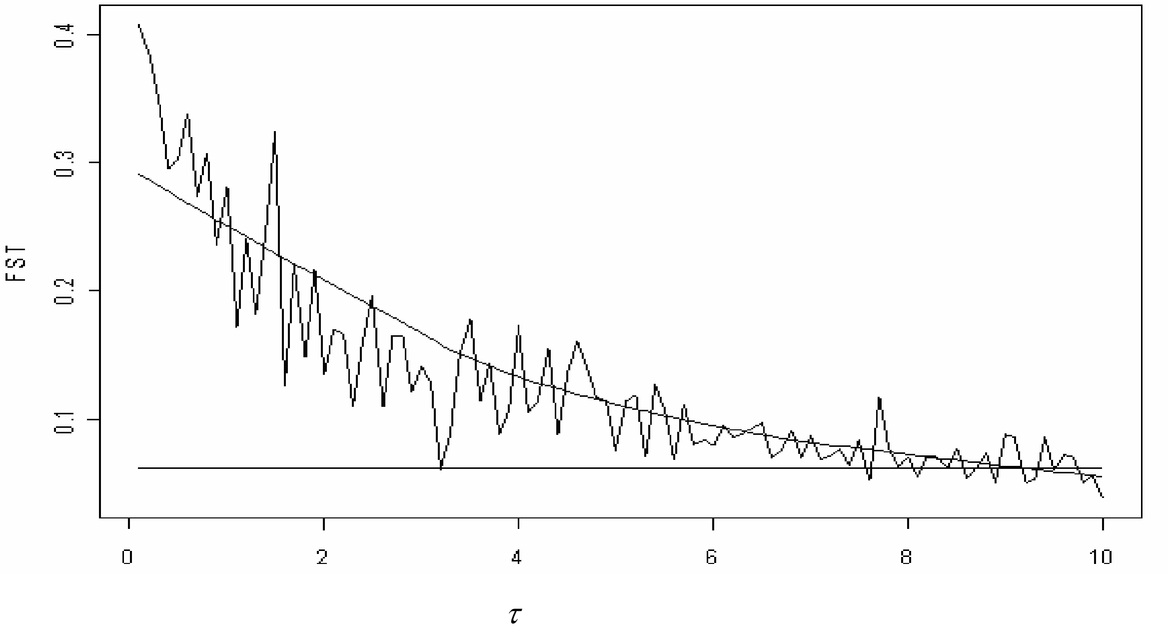


Fig. 3-21 The relationship between Fst and τ , 11 loci were used

The program is **"Fst.R"**

In Fig. 3-21, when Fst=0.062, $\tau > 10$ and this caused that the null hypothesis can not be rejected at 95% significance level, just as showed in Fig. 3-22. The ppoints analysis when $\tau = 1.0$ and $\tau = 10$ are showed in Fig. 3-23 and Fig. 3-24.

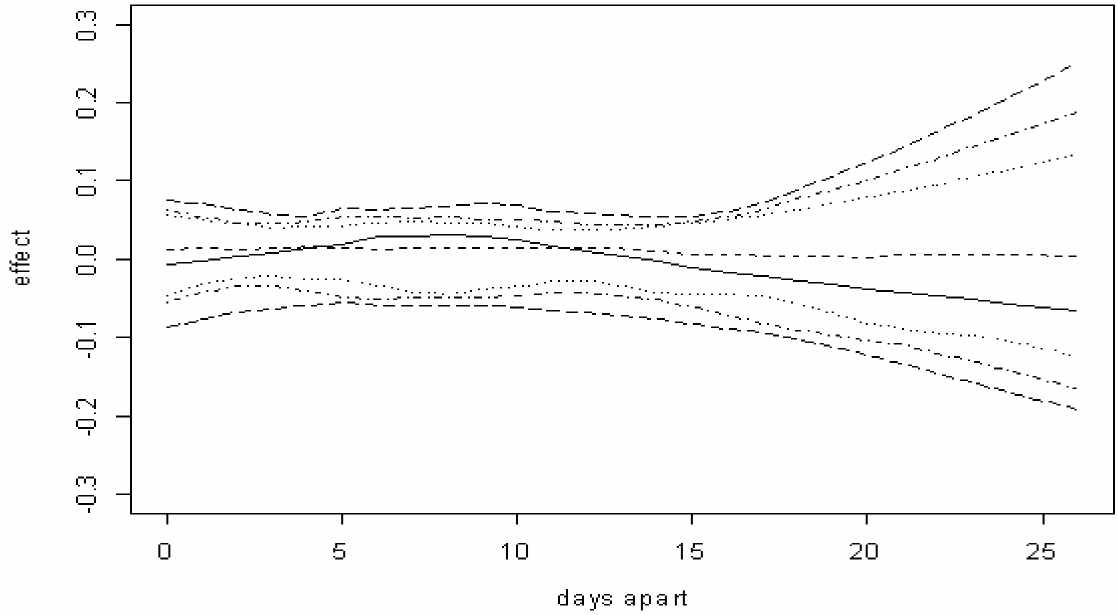


Fig. 3-22 Simulation result with pointwise band when τ =10.0

A probability-probability (P-P) plot is used to see if a given set of data follows some specified distribution. It should be approximately linear if the specified distribution of alternative hypothesis follows the null hypothesis. The probability-probability (P-P) plot is constructed using the theoretical cumulative distribution function, F(x), of the specified model. The values in the sample of data, in order from smallest to largest, are denoted x(1),x(2),.....,x(n). For i = 1, 2, ..., n, F(x(i)) is plotted against [(i – ½)/n]

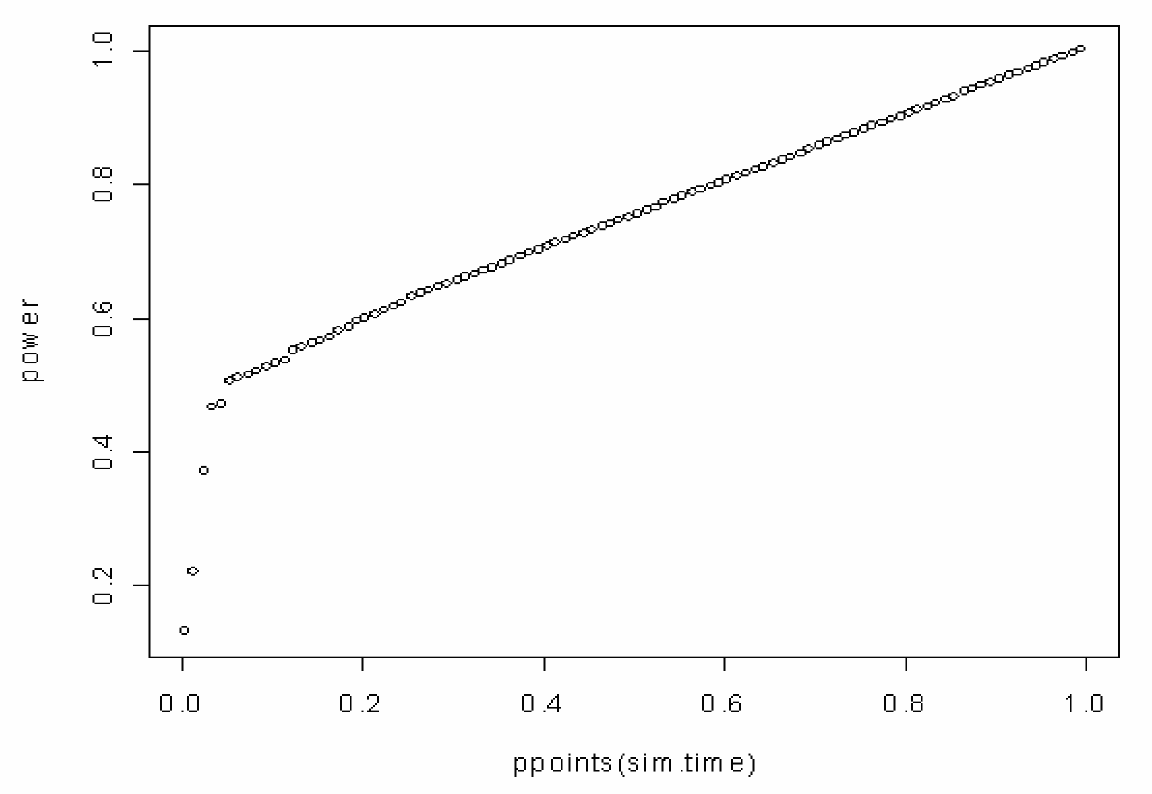


Fig. 3-23 ppoints analysis when $\tau = 1.0$

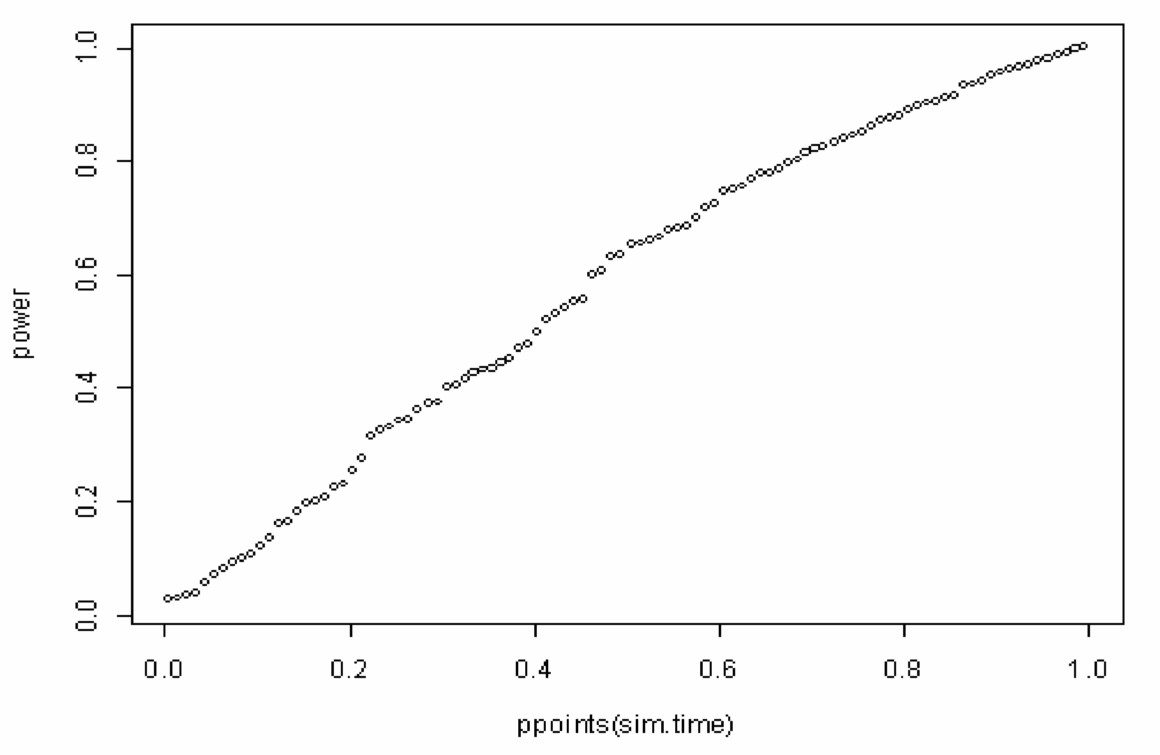


Fig. 3-24 ppoints analysis when $\tau = 10$

The program is “**Whole Programe in R**”

The ppoints figures show that with the increasing of τ , the trend becomes more linear, which means there is less difference between alternative hypothesis and null hypothesis.

Then, a different distribution hypothesis is used to compare its influence on the result. The new hypothesis doesn’t move the samples less than two in a year and the W group is smaller than that in the first hypothesis: $\mu_1 = 293$, $\sigma_1 = 2.5$, $t_1 = 1$; $\mu_2 = 300$, $\sigma_2 = 2.5$, $t_2 = 1$; $\mu_3 = 307$, $\sigma_3 = 2.5$, $t_3 = 1$.

The new distribution hypothesis and the whale samples were drawn in Fig. 3-25 and the results are showed in Fig. 3-26 and Fig. 3-27:

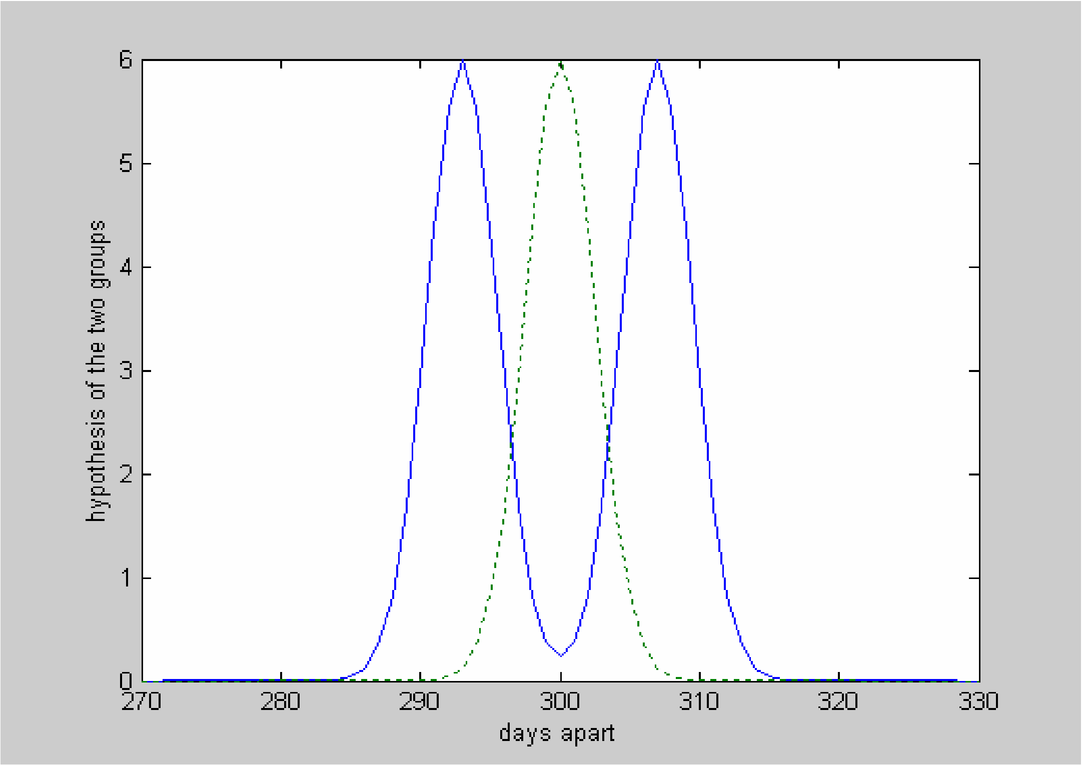


Fig. 3-25 new distribution hypothesis

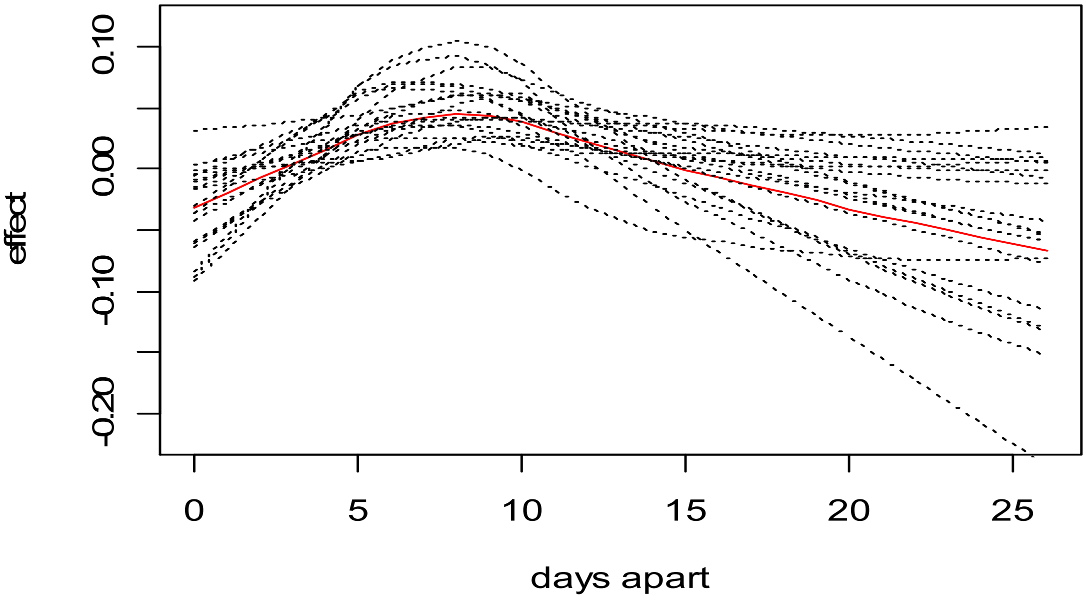


Fig. 3-26 20 simulation results $\tau = 10$

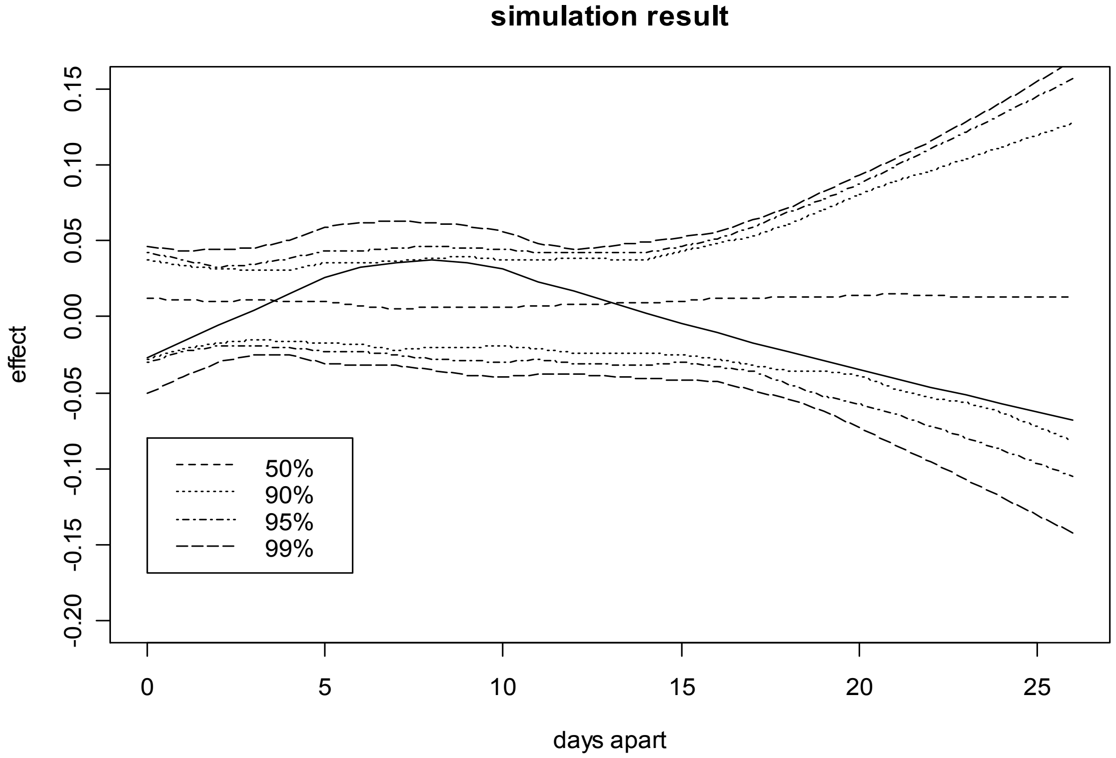


Fig. 3-27 Simulation result $\tau = 10$

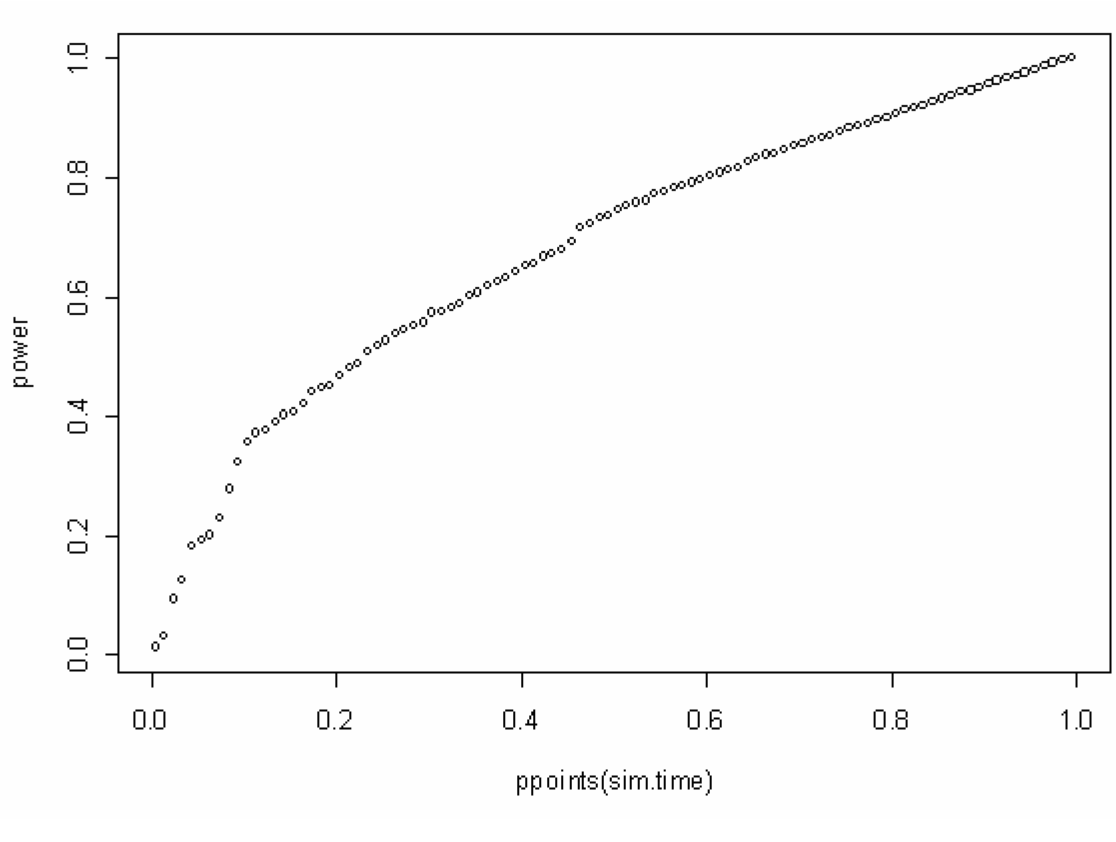


Fig. 3-28 ppoints analysis when $\tau = 10$

Comparing the results of the two different distribution hypothesis, the curve of the second distribution hypothesis is much more significant than that of the first distribution hypothesis.

Summary

A framework for simulating the Chukchi Circuit hypothesis has been developed to investigate whether a plausible variants of the hypothesis is consistent with the observed Oslo bump. To get a bump neatly as marked as that found by Jorde et al (2004) rather extreme parameter values are required in the case I have considered. I therefore suggest that the Chukchi Circuit hypothesis cannot alone explain the observed pattern.

Chapter 4   The Economics of need

How to fulfill people’s need is an important topic in Economics. Needs are defined as goods or services that are required, which include the needs for food, clothing, shelter and health care. Wants are goods or services that are not necessary but we desire or wish for. Sometimes needs and wants overlap. For example, cake is a food, but it is a want, not a need. In general, you need a basic diet to survive, but that diet doesn’t need to include cake. An important part of the economics is the distribution of resources or goods so that people's needs are met. This is especially true in times of scarcity when there are not enough resources, goods or services.

The objective of resources management should be achieving efficiency and at the same time, giving attention to equity. Taking care of posterity must be also set as a moral obligation and sustainability constraints have to be included in the optimality problem.

Answer by resource economics

Resource economics and welfare economics are two branches of economics. They answer the question of how to fulfill people’s need in different ways. Resource economics has evolved as the idea of "natural resources" and "human resources" were challenged by the ideas of "natural capital" and "human capital". It was a major influence on the theory of “Natural Capitalism” and of “eco-villages” [18]. Three themes are emphasized in resource economics about resource management: efficiency, optimality and sustainability.

Economic efficiency is a general term for the value assigned to a situation by some measure designed to capture the amount of waste or "friction" or other undesirable economic features present. There are several measures of economic efficiency such as Pareto efficiency, productive efficiency and distributive efficiency.

However, the economic concept of efficiency is not the only thing that a society might care about. In particular, the theorem says nothing about the distributional equity of the outcome. Economic efficiency means that the "correct people" (those who can afford it) will get the "correct goods and services" (whatever they want). It may allocate resources to people who are the most successful at gaining social power. The rich may get richer and the poor may get poorer. Someone even think "Economic efficiency" means "economic anarchy" (no government regulations). It leads to "Social Darwinism" -- survival of the economically fittest -- the rest can suffer and die.

Economics optimality means a society maximizing its overall objectives under relevant constraints. According to second welfare theorem, resource allocation cannot be optimal without being efficient, which is necessary, but not sufficient.

The society's overall objective may not concern the welfare of its posterity's. In this situation, sustainability cannot be fulfilled. It is especially critical to the nonrenewable resources and the dying off species. Taking care of posterity must be set as a moral obligation and sustainability constraints have to be included in the optimality problem.

According to first welfare theorem, any competitive equilibrium leads to an efficient allocation of resources [20]. This theorem appears to make a case for non-intervention: let the markets do the work and the outcome will be desirable. The theorem is often taken to be an analytical confirmation of Adam Smith's "invisible hand" hypothesis, namely that competitive markets tend toward the efficient allocation of resources.

In realities, it is very difficult to find an equilibrium point between equity and efficiency. Take building bridge for an example. It is much more efficiency to give the contract to a company and let the company to manage it. However, it may be unfair to the people living in the near villages. They are poor and have to pay for every time passing the bridge. This is a problem happening in the developing countries like China. Development leads to unjust. It is believed that along with the development of economics, the market will allocate the resources to be more and more reasonable and the unjust will be lessened.

It is the government's duty to make the decision. Most of the developing countries are anxious to develop their economy and overlook the bad consequences. Gradually, people will realize the development of economy is not the only object and it will not necessarily lead to human happiness.

Answer by welfare economics

Welfare economics is another branch of economics that uses microeconomic techniques to simultaneously determine the allocational efficiency of a macroeconomy and the income distribution consequences associated with it. Welfare economics concerns the welfare of individuals as opposed to groups, communities and societies and supposes the welfare can be adequately measured in monetary terms. There are two sides to welfare economics: economic efficiency and income distribution.

One important measure of efficiency in welfare economics was Abba Lerner's proposed distributive efficiency. Situations are considered to have distributive efficiency when goods are distributed to the people who can gain the most utility from them. Many economists use Pareto efficiency as their efficiency goal. According to this measure of social welfare, a situation is optimal only if no individuals can be made better off without making someone else worse off. This ideal state of affairs can only come about if four criteria are met [20]:

- The marginal rates of substitution in consumption are identical for all consumers. This occurs when no consumer can be made better off without making others worse off.
- The marginal rate of transformation in production is identical for all products. This occurs when it is impossible to increase the production of any good without reducing the production of other goods.
- The marginal resource cost is equal to the marginal revenue product for all production processes. This takes place when marginal physical product of a factor must be the same for all firms producing a good.
- The marginal rates of substitution in consumption are equal to the marginal rates of transformation in production, such as where production processes must match consumer wants.

There are a number of conditions that, most economists agree, may lead to inefficiency. They include:

- Imperfect market structures, such as a monopoly, monopsony, oligopoly, oligopsony, and monopolistic competition.

- Factor allocation inefficiencies in production theory basics.
- Market failures and externalities; there is also social cost.
- Price discrimination and price skimming.
- Long run declining average costs in a natural monopoly.
- Certain types of taxes and tariffs.

To determine whether an activity is moving the economy towards Pareto efficiency, two compensation tests have been developed. Any change usually makes some people better off while making others worse off, so these tests ask what would happen if the winners were to compensate the losers. Using the Kaldor criterion, an activity will contribute to Pareto optimality if the maximum amount the gainers are prepared to pay is greater than the minimum amount that the losers are prepared to accept. Under the Hicks criterion, an activity will contribute to Pareto optimality if the maximum amount the losers are prepared to offer to the gainers in order to prevent the change is less than the minimum amount the gainers are prepared to accept as a bribe to forgo the change. The Hicks compensation test is from the losers' point of view, while the Kaldor compensation test is from the gainers' point of view. If both conditions are satisfied, both gainers and losers will agree that the proposed activity will move the economy toward Pareto optimality. This is referred to as Kaldor-Hicks efficiency or the Scitovsky criterion [16].

A Pareto efficiency does not necessarily mean equity. Suppose a society with two men, a rich and a poor. The rich owns all and the poor has nothing. This is the Pareto efficiency of this society because we can not make the poor better off without making the rich worse off except that we can move the production frontier outward.

The basic welfare economics problem is to find the theoretical maximum of a social welfare function, subject to various constraints such as the state of technology in production, available natural resources, national infrastructure, and behavioral constraints such as consumer utility maximization and producer profit maximization. To attain equity, the utility of the poor is seen to be greater value than that of the rich when a social welfare function is built by summing up the utility of each individual.

Theoretically, the optimal solution of the welfare function can be solved but in realities, it is almost impossible to solve a question like that because there are millions of consumers and thousands of producers, each of whom has his own utility function and production function. Even if you can find a solution to the complicated mathematical function, the result means little to an actual problem.

To sum up, the resources management's objective should be achieving efficiency and at the same time, giving attention to equity, the welfare of this generation and our posterity. The rankings of allocation must be based on ethical criterions. Culture background must be taken into consideration when building utility functions.

For an example, we can study the consumption of whales by Inuit and other people in an Edgeworth box:

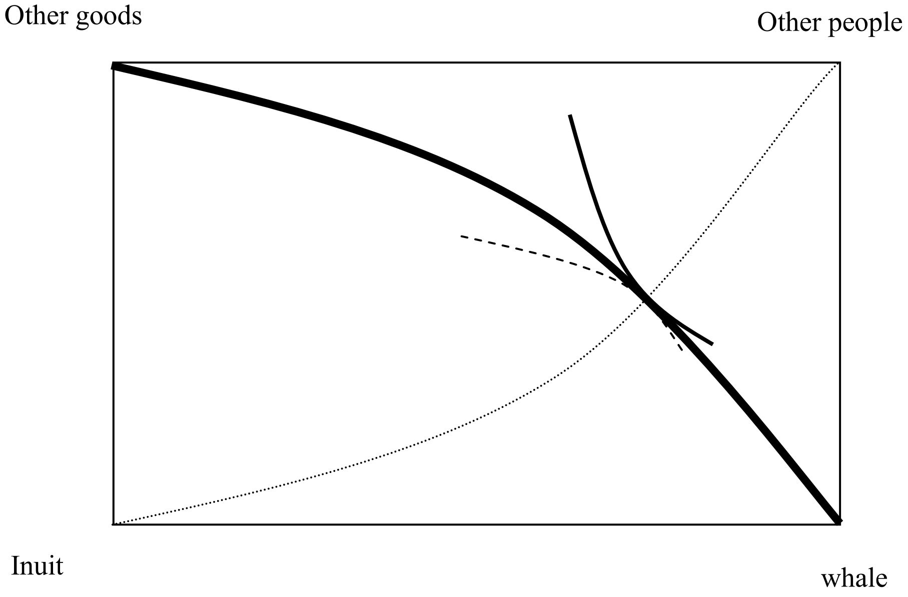


Fig. 2-1 An Edgeworth box

The Pareto efficiency is achieved when the marginal rates of substitution in consumption are identical to the Inuit and other people in this world. The marginal rate of substitution in consumption should also equal to the marginal rate of transformation in production. Take sustainability into consideration, the marginal rate of the present value should be equal to the marginal rate of cost, the Hotelling's rule of a renewable resource.

Tax and quota can be used in resource management [18]. Assuming the utility function is f(x), the tax rate is t, maximize the utility:

$$\text{Max } \{f(x) - tx\}$$

$$\Rightarrow t = f'(x)$$

Which means the tax rate should be equal to the marginal acquisition of the utility of the consumer at amount x.

Inuit is poor and we can not use tax on them. However, we can give them subsidies for not consuming. The subsidy rate is s. To maximize their utility:

$$f(x) - f(x^0) = \text{increased utility from consumption}$$

$$s(x - x^0) \quad = \text{paid subsidy}$$

$$Max\{f(x) - f(x^0) - s(x - x^0)\}$$

$$\Rightarrow s = f'(x)$$

This is the same as the tax rate.

Difficulties

In realities, there are some difficulties in fulfilling people's need. First, it is quite difficult to distinguish needs from wants because the difference between them is so fuzzy. It depends on the people's purchasing power and the cost of production, the marginal utility of money. Different people have different standard. To the poor, they may have to live in a 5 m^2 room and eat bread all day. But to a millionaire, he may need big houses and a dozen cooks to make meal for him and his friends. What's more, the standard varies along with the time and place. In a rich harvest year, almost everyone can have barely good food. But in a bad harvest year, someone has to suffer hunger. In Europe, there is plenty of water, but in African desert, a bath can be a luxury.

Second, it is difficult to value the things in this world. It fully depends on how people need it, in another world, how much people would like to pay for it. What is the price of a beautiful sunset, a clean blue crystal sea, a song of a happy bird in the woods, the life of a dignity whale? The whale may be very important to the Inuit but means nothing to a businessman in Beijing. This will influence the environment policy since most of the people do not concern the things far away.

Third, the economics tries to constrain its subject to be resources allocation, production, distribution and consumption. This is to avoid infringing on other disciplines, such as politics and sociology. Economists are trained to believe that "money" has nothing to do with politics and is simply a medium of exchange. But even the casual observer can see that money is a social power because it empowers people to buy and do the things they want -- including buying and doing other people: politics. In fact, economists do appeal to bring together all kinds of people, thinkers, activists, academics and policy makers, businessmen, economists and campaigners, writers and opinion formers to work for their ideal, which is politics.

Fourth, the human society is perhaps the most complicated organization we have ever met. Everyone has his own sentiment, thought, and beliefs. What’s more, there are customs, traditions, cultures and religions around us. We cannot evaluate them precisely and we don’t know how these things will affect us. The bowhead whales, for an example, should certainly be protected as an endangered and dignity life. However, we must consider the need of the Inuit’s because it is not only a food resource of them, but also a culture.

Chapter 5 The simulation on bowhead whale management

In the following, I will do some simulation based on Pella-Tomlinson model. First, if the growth parameter r is a constant, we can find the maximum we can catch per year. If the growth parameter varies, the risk of extinction goes up and I will try to estimate the probability at different number of harvest.

Assuming the stock of the whales follows Pella-Tomlinson model, and there is a random noise in r: $r(t) = r_0 + r'$. According to IWC, r_0 is about 0.03 and let $r' \sim N(0, 0.04)$. The simulation shows that the population growth will change greatly.

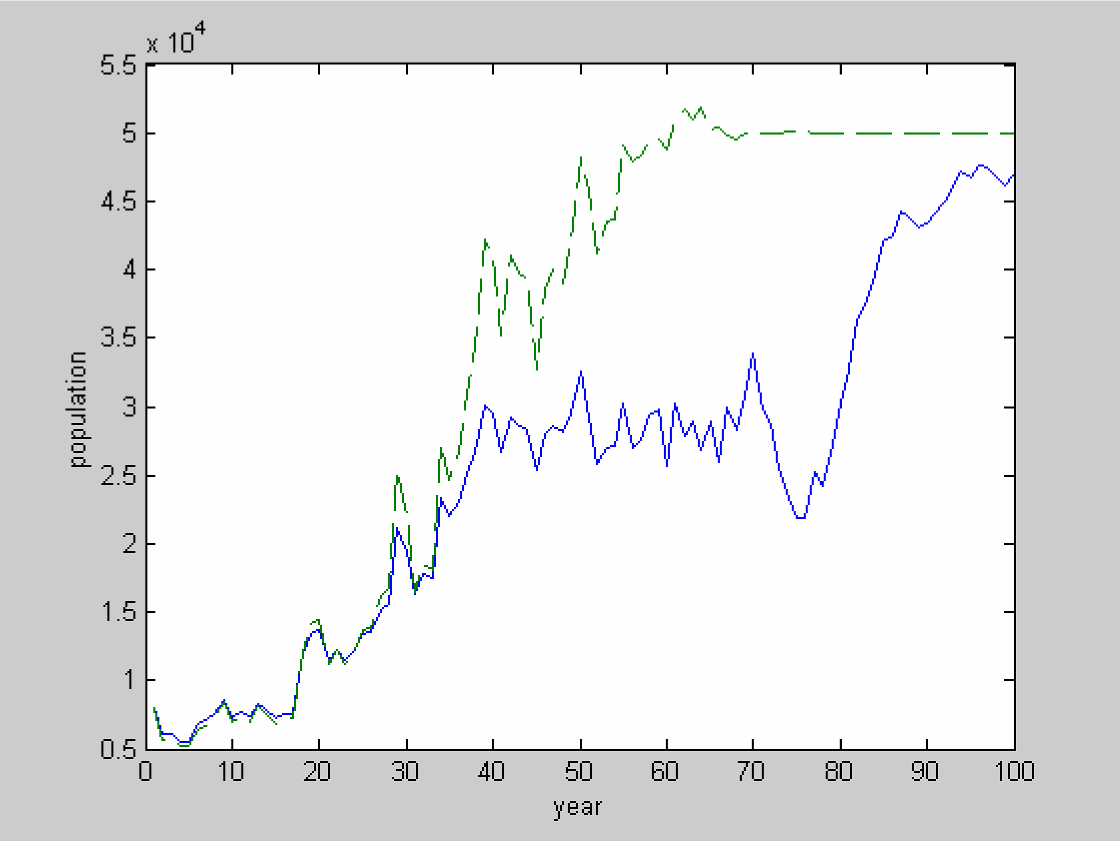


Fig. 5-1 The Pella-Tomlinson model with noise in r, without consumption

(K = 20000, X_0 = 1000, r = 3%, $r' \sim N(0, 0.04)$ $z_1 = 1$, $z_2 = 5$)

The program is “**Fish Management: Pella_Tomlinson_seq.m**”, If we let the Inuit to consume a fix amount whales in the future, what will happen when r is a constant r and a variable.

$$Y_{t+1} = X_t + rX_t(1-(\frac{X_t}{K})^z) \text{ and } Y_{t+1} = X_t + C_t$$

First, we can find the maximum number C(t) which will make the whale stock constant in the future if r is a constant:

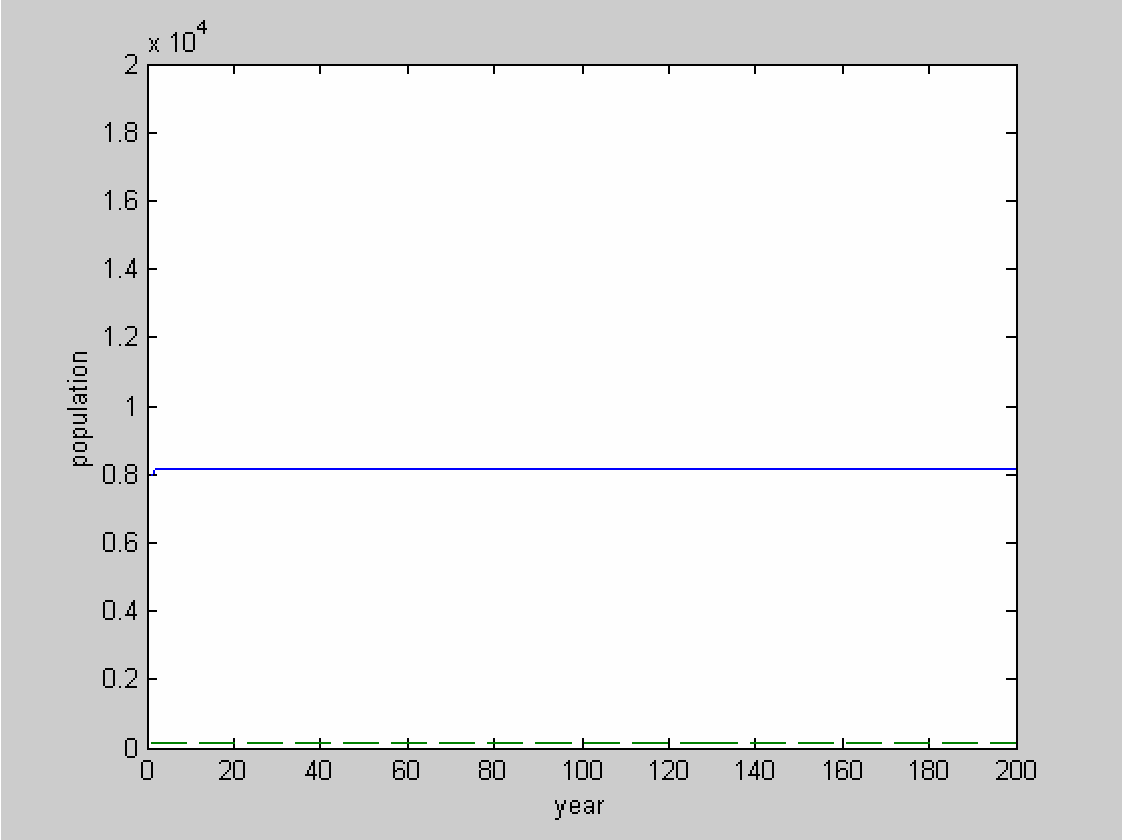


Fig. 5-2 The simulation result when r is a constant

(K = 20000, X_0 = 8000, r = 3%, z_1 = 1, C=144)

The program is “**Fish Management: constant and variable.m**”

The simulation result shows that 144 whales can be caught each year to keep the stock constant in the future. However if there is noise in r, the risk of extinction will be higher:

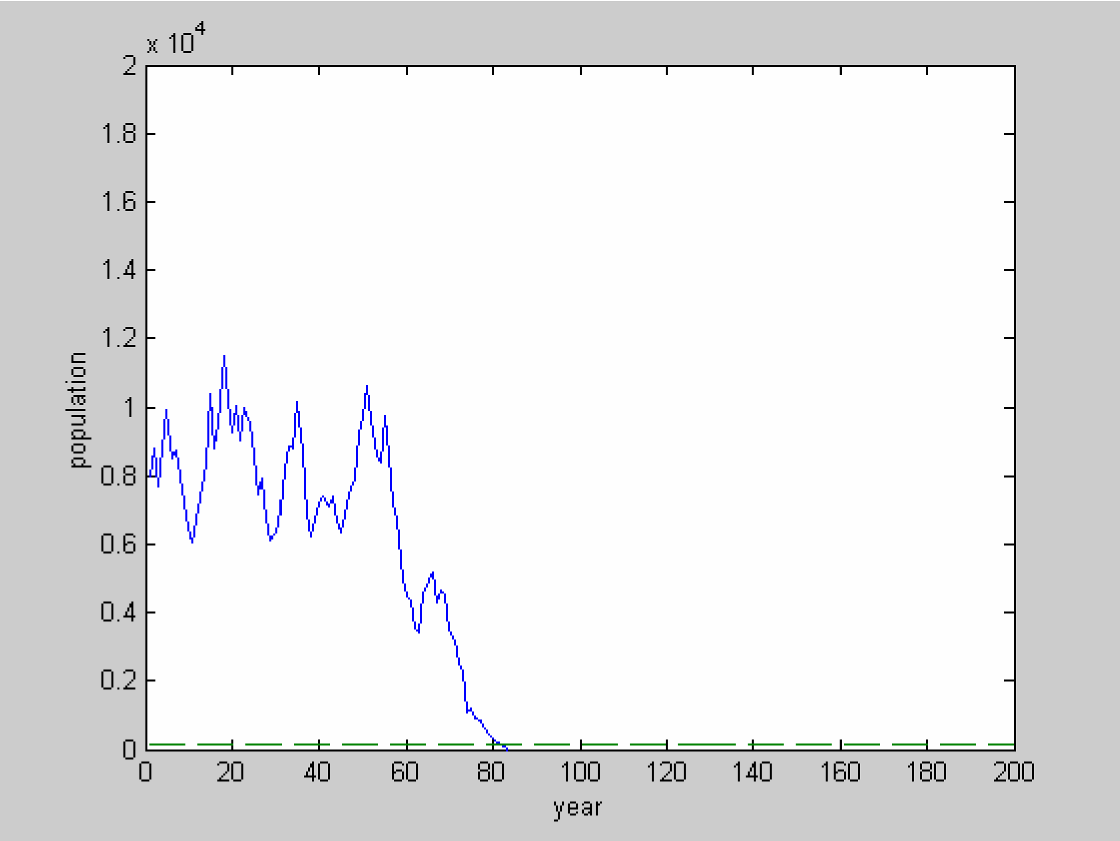


Fig. 5-3 The simulation result when there is a noise in r

(K = 20000, $X_0 = 8000$, r = 3%, $z_1 = 1$, C=120)

The program is “**Fish Management: constant and variable m**”

Fig. 5-3 shows when there is a noise in r, the species will die off after 80 years if we catch 120 whales per year. We have to catch less per year to prevent this happen which is showed in Fig. 5-4 as the catch number decrease to 100 per year.

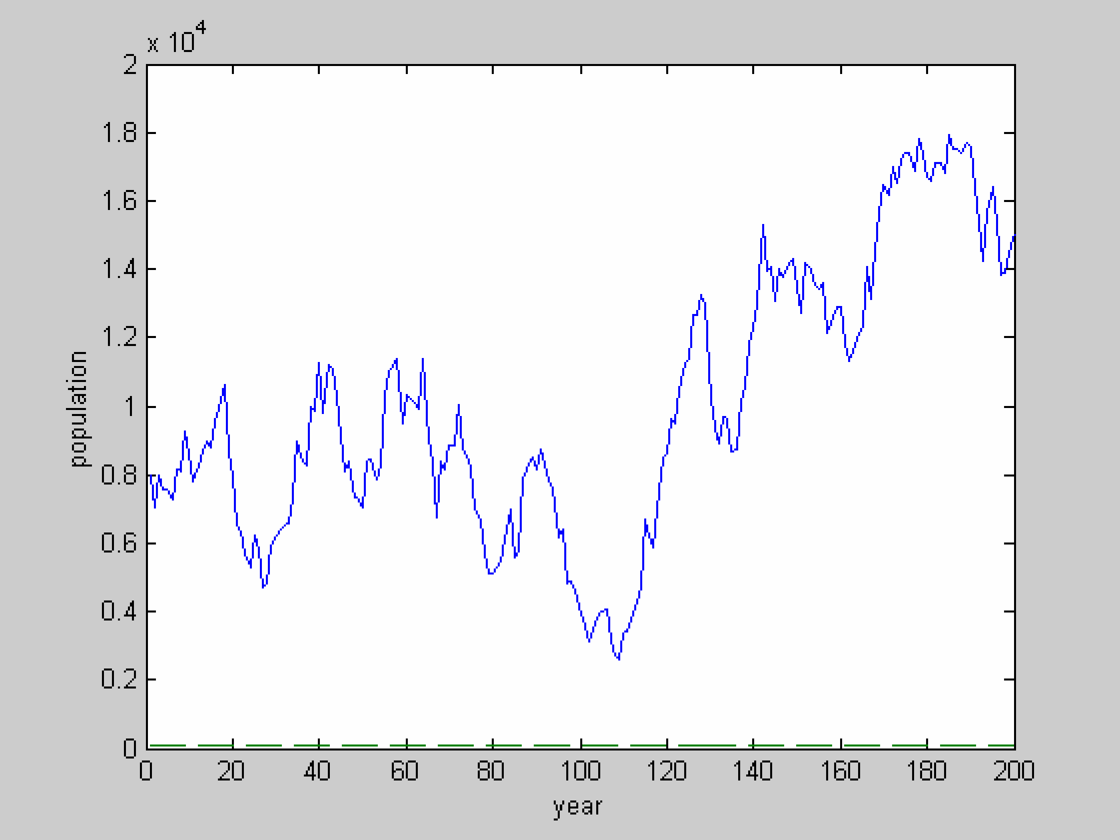


Fig. 5-4 The simulation result when there is a noise in r

(K = 20000, X_0 = 8000, r = 3%, z_1 = 1, C=100)

The program is "**Fish Management: constant and variable m**"

However, both of the figures are the random results of one-time simulation. They can't be used to determine the risk of extinction. To study the risk of extinction, I simulated 50 times and calculate the frequency of extinction at each catch number.

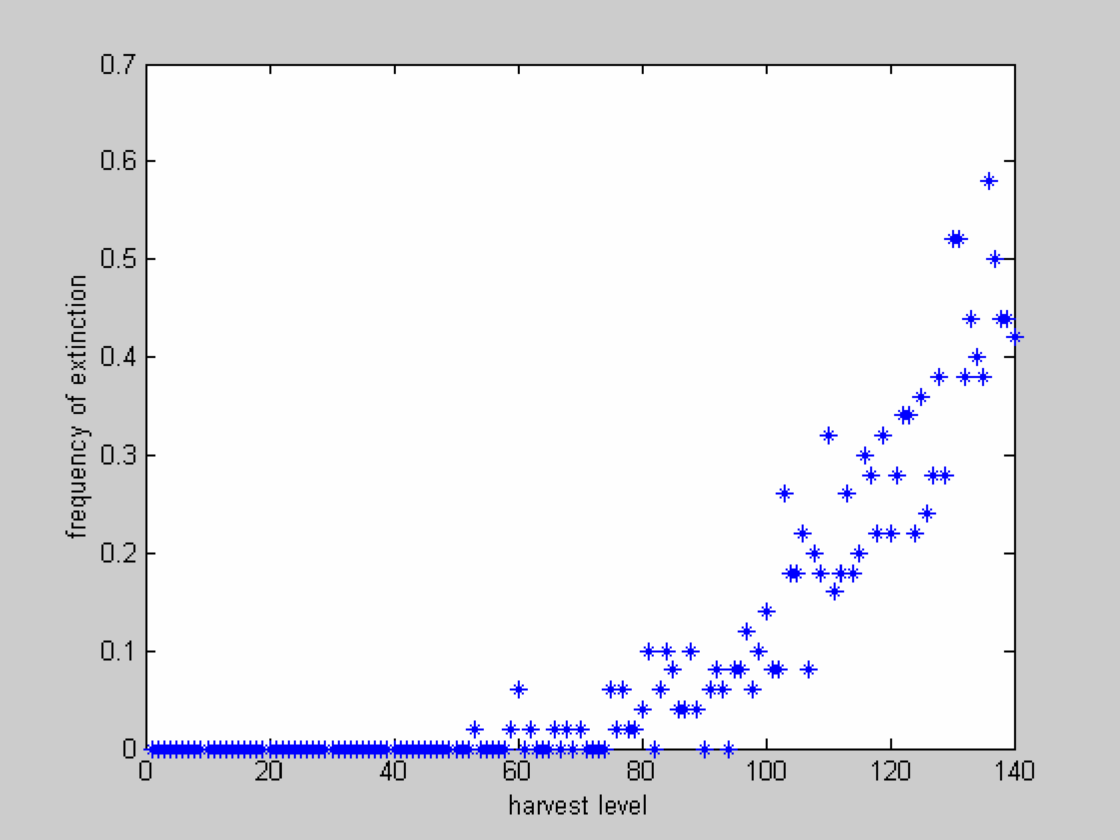


Fig. 5-5 The relative frequency of extinction

By constant harvest level

(K = 20000, X_0 = 8000, r = 3%, $r' \sim N(0,0.01)$, z_1 = 1,

n=50 simulation for each constant level of catch)

Regress the probability (Y) on the catch number (X):

$Y = 0$, for $X <= 76$,

$Y = \beta_1 X + \beta_0$, for $X > 76$

$\Rightarrow \quad \beta_1 = 0.006416, \quad \beta_0 = -0.488342$

Which means it is almost safe to catch less than 80 whales per year. When the number is more than 80, the risk of extinction will increase 0.6% if one more whale is caught.

If there is a bigger noise in r, then a higher risk of extinction will be. The Fig. 5-6 is the simulation result when $r' \sim N(0,0.04)$.

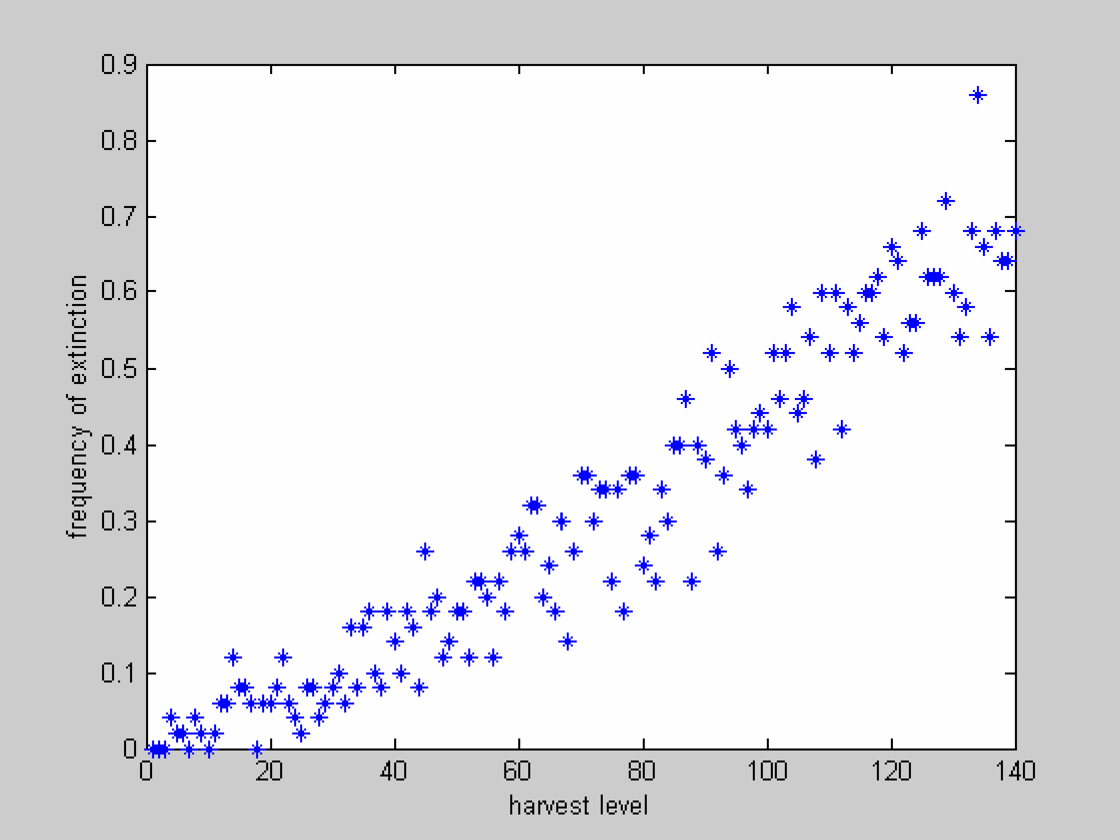


Fig. 5-6 The relative frequency of extinction

By constant harvest level

(K = 20000, X_0 = 8000, r = 3%, $r' \sim N(0,0.04)$, z_1 = 1,

n=50 simulation for each constant level of catch)

Regress the probability (Y) on the catch number (X) and we can get:

$\beta_1 = 0.004925$, $\beta_0 = -0.045059$

Which means it is not safe to catch any whales. The risk of extinction will increase 0.49% if one more whale is caught.

On the other hand, if the stock is smaller, the risk of extinction will go up with the same harvest level. The following figure shows if there are only 4000 whales in stock, the safe catch limit should be less than 40 per year.

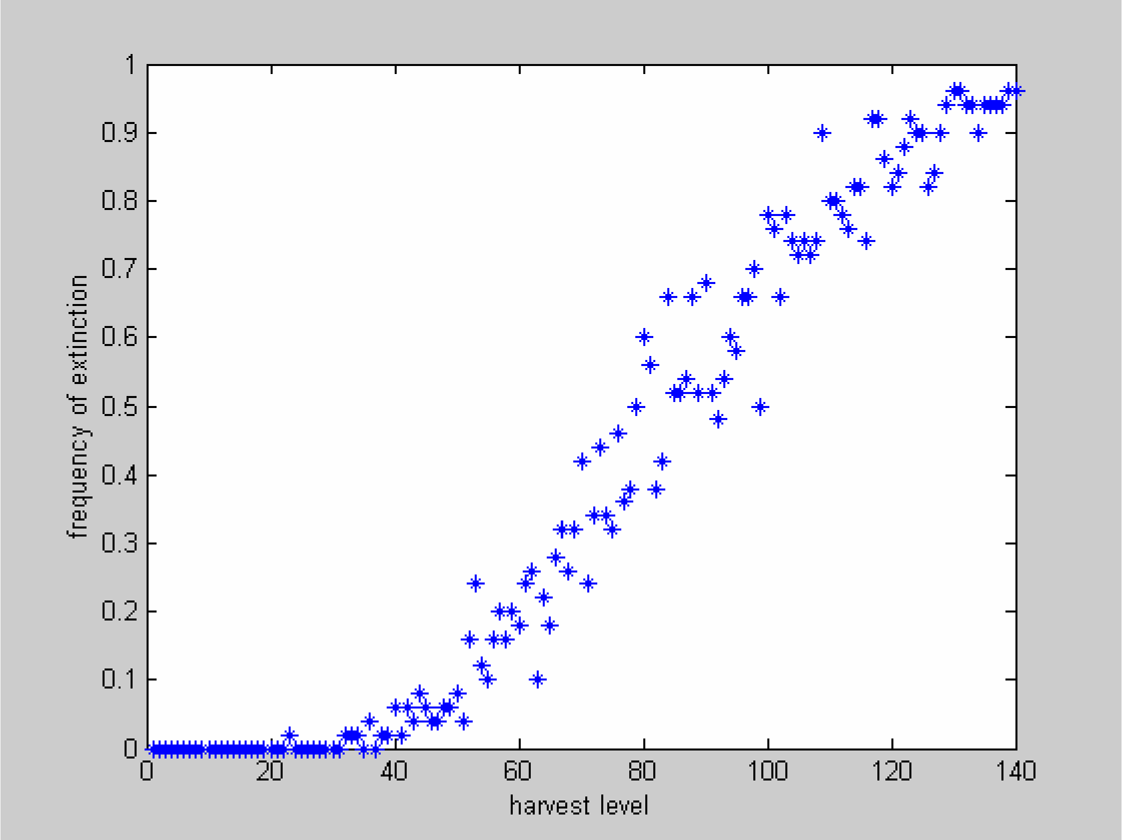


Fig. 5-7 The relative frequency of extinction

By constant harvest level

(K = 20000, $X_0 = 4000$, r = 3%, $r' \sim N(0,0.01)$, $z_1 = 1$,

n=50 simulation for each constant level of catch)

The program is "**Fish Management: statistics.m**"

To sum up, the risk of extinction depends on the variability of the growth rate r and the population of stock. What will influence the variability of r? There are many factors, such as the chance of a male meeting a female, the chance of pregnancy, the pollution and the change of climate may cause the food supply decline and so on. These unknown factors will make it difficult to estimate r.

The population of the stock is another determining factor. If the stock is proved to be a smaller one, a more conservative policy should be implemented.

Conclusion

Constant harvest rate seems dangerous. We really don’t know the stochastic element in the population dynamics.

The adaptive procedure developed by the AWMP group seems reasonable. It sets the quota equal to the stated need (the quota required by the Inuits) if abundance estimates catch series and further biological information indicates that this catch limit is sustainable in the long run. If however the rather complicated calculations indicate this not to be the case, a lower catch limit is found that seems sustainable.

If the Chukchi Circuit hypothesis can be proved to be true, then the catch limit has to be reconsidered in order to prevent the extinction of a small sub stock.

Appendix Programs

```
% Ch3
% Whale_DataPro_Main.m (m is the suffix of programs in Matlab)
clear;

global data3;
global E_hyp_p;
global W_hyp_p;

%data of observe and with same media
data3=whale_observe;
%aggregate whale data
data_agg=aggregate(data3);
%hypothesis probability
mu_1=293;
mu_2=300;
mu_3=306;
delta_1=3.5;
delta_2=3;
delta_3=5;
R1=0.44;
R2=0.41;
R3=1;
[E_hyp_p,W_hyp_p]=hypo_prob(mu_1,mu_2,mu_3,delta_1,delta_2,delta_3,R1,R2,R3);
for i=1:10;
    j=1;
    while data3(i,j)>0;
        P_W=W_hyp_p(data3(i,j));
        Prob_whale=rand(1);
        if Prob_whale>P_W
            Whale_E(i,j)=1;
            Whale_W(i,j)=0;
        else
            Whale_E(i,j)=0;
            Whale_W(i,j)=1;
        end
        j=j+1;
    end
end
figure(6);
plot_E_W(Whale_E,Whale_W,data3);
hold off
```

% whale_observe.m

```
function data3=whale_observe()
data2=[303   304   305   305   306   312   312   318   319   320   322   322   323   324   325   326   327   0
       311   314   314   0     0     0     0     0     0     0     0     0     0     0     0     0     0     0
       297   299   300   306   0     0     0     0     0     0     0     0     0     0     0     0     0     0
       318   0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0
       300   0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0
       284   287   290   291   292   293   0     0     0     0     0     0     0     0     0     0     0     0
       281   282   285   285   285   285   288   290   290   293   295   295   299   0     0     0     0     0
       277   326   0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0
       276   277   277   289   0     0     0     0     0     0     0     0     0     0     0     0     0     0
       306   314   0     0     0     0     0     0     0     0     0     0     0     0     0     0     0     0];

figure(1);
plot_whale(data2);
hold off;
%same median time
for i=1:10;
    k=0;
    j=1;
    while data2(i,j)>0;
        k=k+1;
        j=j+1;
    end
    medial_data(i)=floor((data2(i,1)+data2(i,k))/2);
    for j=1:k;
        data3(i,j)=data2(i,j)+(300-medial_data(i));
    end
    data3(i,k+1)=0;
end
```

% plot_whale.m

```
function plot_whale(day);
for i=1:8;
    j=1;
    while day(i,j)>0;
        Tu_1_i(j)=day(i,j);
        Tu_1_j(j)=2003-i;
        j=j+1;
    end
    plot(Tu_1_i,Tu_1_j,'o');
    hold on;
end
```

```
for i=9:10;
        j=1;
        while day(i,j)>0;
                Tu_1_i(j)=day(i,j);
                Tu_1_j(j)=1992-(i-9)*2;
                j=j+1;
        end
        plot(Tu_1_i,Tu_1_j,'o');
end
```

% aggregate.m

```
function data_agg=aggregate(data3)
k=1;
for i=1:10;
        for j=1:17;
                if data3(i,j)>0;
                        data_4(k)=data3(i,j);
                        k=k+1;
                end
        end
end
data_agg=sort(data_4);
data_agg(54)=0;
%give number according to day
i=1;
j=2;
k=1;
while data_agg(i)>0;
        data_num(k)=data_agg(i);
        num_day(k)=1;
        while data_agg(i)==data_agg(j);
                num_day(k)=num_day(k)+1;
                j=j+1;
        end
        i=j;
        j=j+1;
        k=k+1;
end
figure(3);
plot(data_num,num_day,'o');
axis([270,330,0,10]);
```

% hypo_prob.m

```
function [E_hyp_p,W_hyp_p]=hypo_prob(mu_1,mu_2,mu_3,delta_1,delta_2)
for i=270:330
    E_hyp_1(i)=hypo_distribution(i,mu_1,delta_1);
    E_hyp_2(i)=hypo_distribution(i,mu_3,delta_1);
    W_hyp_1(i)=hypo_distribution(i,mu_2,delta_2);
end
E_hyp=E_hyp_1+E_hyp_2;

figure(4);
i=1:330;
plot(i,E_hyp,'-',i,W_hyp_1,':');
axis([270,330,0,0.1]);
set(gca,'xtick',[270,280,290,300,310,320,330]);
%hyp_probability
for i=270:330;
    E_hyp_p(i)=E_hyp(i)/(E_hyp(i)+W_hyp_1(i));
    W_hyp_p(i)=W_hyp_1(i)/(E_hyp(i)+W_hyp_1(i));
end
figure(5);
i=1:330;
plot(i,E_hyp_p,'-',i,W_hyp_p,':');
axis([270,330,0,1]);
set(gca,'xtick',[270,280,290,300,310,320,330]);
```

% hypo_distribution.m

```
function y=hypo_distribution(i,mu,delta)
s=(i-mu)^2;
t=2*delta^2;
u=1/sqrt(t*pi);
y=u*exp(-s/t);
```

% plot_E_W

```
function plot_E_W(Whale_E,Whale_W,data3);
for i=1:8;
    for j=1:17;
        if Whale_E(i,j)>0;
            x1(j)=data3(i,j);
            y1(j)=2003-i;
            plot(x1(j),y1(j),'o');
            hold on;
        end
    end
end
```

```
for i=1:8;
    for j=1:17;
        if Whale_W(i,j)>0;
            x2(j)=data3(i,j);
            y2(j)=2003-i;
            plot(x2(j),y2(j),'*');
            hold on;
        end
    end
end
for i=9:10;
    for j=1:17;
        if Whale_E(i,j)>0;
            x3(j)=data3(i,j);
            y3(j)=1992-(i-9)*2;
            plot(x3(j),y3(j),'o');
            hold on;
        end
    end
end
for i=9:10;
    for j=1:17;
        if Whale_W(i,j)>0;
            x4(j)=data3(i,j);
            y4(j)=1992-(i-9)*2;
            plot(x4(j),y4(j),'*');
            hold on;
        end
    end
end


% Ch3
% Whale_GenePro_Main.m
clear;
Agg_Aij(1:245,1:2)=0;
Agg_Aij_1(1:245,1:2)=0;

global data3;
global E_hyp_p;
global W_hyp_p;

for time=1:100;
```

```
for i=1:10;
    j=1;
    while data3(i,j)>0;
        P_W=W_hyp_p(data3(i,j));
        Prob_whale=rand(1);
        if Prob_whale>P_W
            Whale_E(i,j)=1;
            Whale_W(i,j)=0;
        else
            Whale_E(i,j)=0;
            Whale_W(i,j)=1;
        end
        j=j+1;
    end
end

%generate E or W according to hypothesis
tau=2.0;
[Pi_e,Pi_w,Pi_prob_e,Pi_prob_w,Pi_e_num,Pi_w_num,alpha]=gene_pool(tau);

for i=1:10;
    for j=1:17;
        if Whale_E(i,j)==1
            Whale_num(i,j,:,:)=Whale_gene_e(Pi_prob_e,Pi_e_num);
        else if Whale_W(i,j)==1
                Whale_num(i,j,:,:)=Whale_gene_w(Pi_prob_w,Pi_w_num);
            end
        end
    end
end

%calculate h
h=homozygosity(Whale_num,data3);

%Compare and calculate I
I=Compare(Whale_num,data3);

%Calculate Aij
[Aij,Dij,Ai_j]=Calculate_Aij(h,I,data3);

%Ai_j(m,1)=Aij(i,p,j)
%Ai_j(m,2)=Dij(i,p,j)
A_i_j=sortrows(Ai_j,2);
Agg_Aij(:,time)=A_i_j(:,1);
```

```
end
Agg_Aij(:,time+1)=A_i_j(:,2);
save my_data_20_1.out Agg_Aij -ASCII
figure(7);
plot(A_i_j(:,2),A_i_j(:,1),'*');
```

% gene_pool.m

```
function [Pi_e,Pi_w,Pi_prob_e,Pi_prob_w,Pi_e_num,Pi_w_num,alpha]=gene_pool(tau)

global allele_number;
global alpha_t;
alpha_t=Alpha_Proc;

 for i=1:11;
     for j=1:allele_number(i);
        alpha(i,j)=tau*alpha_t(i,j);%alpha after uniform
        Y_gamma_e(i,j)=gamrnd(alpha(i,j),1);
        Y_gamma_w(i,j)=gamrnd(alpha(i,j),1);
    end
    for j=1:allele_number(i);
        Pi_e(i,j)=Y_gamma_e(i,j)/sum(Y_gamma_e(i,:)); %probability of a dirichlet distribution
        Pi_w(i,j)=Y_gamma_w(i,j)/sum(Y_gamma_w(i,:));
        Pi_e_num(i,j)=j;
        Pi_w_num(i,j)=j;
    end

    Pi_e_1(i,:)=Pi_e(i,:);
    for j=1:allele_number(i);
        Min_a=0;
        Min_num=0;
        for k=j+1:allele_number(i);           %Sequence, in case of calculation error
            if Pi_e_1(i,j)>Pi_e_1(i,k);
                Min_a=Pi_e_1(i,j);
                Pi_e_1(i,j)=Pi_e_1(i,k);
                Pi_e_1(i,k)=Min_a;
                Min_num=Pi_e_num(i,j);
                Pi_e_num(i,j)=Pi_e_num(i,k);
                Pi_e_num(i,k)=Min_num;
            end
        end
    end
```

```
Pi_w_1(i,:)=Pi_w(i,:);
for j=1:allele_number(i);
    Min_a=0;
    Min_num=0;
    for k=j+1:allele_number(i);
        if Pi_w_1(i,j)>Pi_w_1(i,k);
            Min_a=Pi_w_1(i,j);
            Pi_w_1(i,j)=Pi_w_1(i,k);
            Pi_w_1(i,k)=Min_a;
            Min_num=Pi_w_num(i,j);
            Pi_w_num(i,j)=Pi_w_num(i,k);
            Pi_w_num(i,k)=Min_num;
        end
    end
end

Pi_prob_e(i,1)=Pi_e_1(i,1);
for j=1:allele_number(i)-1;
    Pi_prob_e(i,j+1)=Pi_prob_e(i,j)+Pi_e_1(i,j+1);
end
Pi_prob_w(i,1)=Pi_w_1(i,1);
for j=1:allele_number(i)-1;
    Pi_prob_w(i,j+1)=Pi_prob_w(i,j)+Pi_w_1(i,j+1);
end
end
```

% Alpha_Proc.m

```
function alpha_t=Alpha_Proc()

data=[147   155   157   159   161   163   165   167   171   181   183   185   187   189
      1     9     3     2     103   10    81     2     1     5     3     2     3    5

      239   241   243   245   247    0     0     0     0     0     0     0     0     0
      6     30    60    126   6      0     0     0     0     0     0     0     0     0

      295   299   301   303   305   307   309    0     0     0     0     0     0     0
      5     93    43     7     79    4     3     0     0     0     0     0     0     0

      193   195   197   199   201   203   205   207   209   211    0     0     0     0
      54     2    19    68     8    31    20    10     5     3     0     0     0     0

      172   174   176   178   180   182    0     0     0     0     0     0     0     0
      3     34    47    64    48    18     0     0     0     0     0     0     0     0
```

```
   156   164   166   168   170   172    0    0    0    0    0    0    0    0
   117     8     7    57    44     1    0    0    0    0    0    0    0    0

   115   130   158   162   166   170   173   174   178   182    0    0    0    0
   50    40    14    31    21    23     7    33    14     1     0    0    0    0

   135   137   139   141   143   147    0    0    0    0    0    0    0    0
   55    26    15    35    83     6     0    0    0    0    0    0    0    0

   93   95   97    99   101   103   105   107    0    0    0    0    0    0
   7    14   126   4    58    22      2     1     0    0    0    0    0    0

   140   142   144   146   148   150   152   154   156   160    0    0    0    0
   1       1    42    42    42    37    38    13     6     2     0    0    0    0

   184   186   190   192    0    0    0    0    0    0    0    0    0    0
   7     160   57    10     0    0    0    0    0    0    0    0    0    0];

  global allele_number;
  global alpha_t;

  for i=1:11;
      frequency(i,:)=data(2*i,:);
      allele_number(i)=0;
      for j=1:14;
          if frequency(i,j)>0;
              allele_number(i)=allele_number(i)+1;%allele's number at each locus
          end
      end
      for j=1:allele_number(i);
          alpha_t(i,j)=frequency(i,j)/sum(frequency(i,:));
      end
end
```

% Whale_gene_e.m

```
function E_num=Whale_gene_e(Pi_prob_e,Pi_e_num)
for i=1:11;
    for j=1:2;
        E_prob(i,j)=rand(1);
        k=1;
        while E_prob(i,j) > Pi_prob_e(i,k)
            k=k+1;
```

```
        end
        E_num(i,j)=Pi_e_num(i,k);
    end
end
```

% Whale_gene_w.m

```
function W_num=Whale_gene_w(Pi_prob_w,Pi_w_num)
for i=1:11;
    for j=1:2;
        W_prob(i,j)=rand(1);
        k=1;
        while W_prob(i,j) > Pi_prob_w(i,k)
            k=k+1;
        end
        W_num(i,j)=Pi_w_num(i,k);
    end
end
```

% homozygosity.m

```
function h=homozygosity(Whale_num,data3)
for q=1:11;
m=1;
    for i=1:10;
        for j=1:17;
            if data3(i,j)>0;
                for k=1:2;
                    Compare_1(m)=Whale_num(i,j,q,k);
                    m=m+1;
                end
            end
        end
    end
    Compare_1(109)=0;
    p=109;
    Dif_alle(q)=0;
    while Compare_1(1)>0
        k=1;
        Agg=1;
        Dif_alle(q)=Dif_alle(q)+1;%the number of different alle of the whale at the same loci
        for j=2:p;
            if Compare_1(1)==Compare_1(j);
                Agg=Agg+1;
```

```
            else
                Compare_2(k)=Compare_1(j);
                k=k+1;
            end
        end
        Pi_num(q,Dif_alle(q))=Agg/108;
        p=p-Agg;
        Compare_1=Compare_2;
        Compare_2(1:k-1)=0;
    end
    h(q)=norm(Pi_num(q,:))^2;
end
```

% Compare.m

```
function I=Compare(Whale_num)
for m=1:52;
    for n=m+1:53;
        for j=1:11;
            Same_alle(m,n,j)=0;
            for k_m=1:2;
                for k_n=1:2;
                    if Whale_num(m,j,k_m)==Whale_num(n,j,k_n)
                        Same_alle(m,n,j)=Same_alle(m,n,j)+1;
                    end
                end
            end
        end
    end
end
I=Same_alle/4;
```

% Ch3

% Fst.R (R is the suffix of programs in R)

```
#Loci=loci and allele data
#nAllel=number of allels present in a locus
#Gen_frequency=allel frequencies within loci
#e_gamma=the gamma random of e-group
#w_gamma=the gamma random of w-group
#e_dirichlet=the dirichlet distribution of e-group
#w_dirichlet=the dirichlet distribution of w-group
#Input the Gen data
Loci<-matrix(scan("../Loci.txt"), nrow=22, ncol=14, byrow=TRUE)
```

```
#Calculate the Gen_frequency
temp_frequency=matrix(0,11,14)
Gen_frequency=matrix(0,11,14)
e_gamma=matrix(0,11,14)
w_gamma=matrix(0,11,14)
e_dirichlet=matrix(0,11,14)
w_dirichlet=matrix(0,11,14)
e_diri_pool=matrix(0,11,14)
w_diri_pool=matrix(0,11,14)
nAllel=c(1:11)
for (i in 1:11){
        temp_frequency[i,]<-Loci[2*i,]
        nAllel[i]=0
        for (j in 1:14){
                if (temp_frequency[i,j]>0){
                        nAllel[i]=nAllel[i]+1
                }
        }
}
Sample_num=50
t_range=100
FST=matrix(0,t_range,1)
for (x in 1:t_range){
t_x=0.1*x

for (i in 1:11){
        for (j in 1:nAllel[i]){
                Gen_frequency[i,j]=temp_frequency[i,j]/sum(temp_frequency[i,])
                e_gamma[i,j]=rgamma(1,shape=t_x*Gen_frequency[i,j])
                w_gamma[i,j]=rgamma(1,shape=t_x*Gen_frequency[i,j])
        }
        for (j in 1:nAllel[i]){
                e_dirichlet[i,j]=e_gamma[i,j]/sum(e_gamma[i,])
                w_dirichlet[i,j]=w_gamma[i,j]/sum(w_gamma[i,])
                }
        e_diri_pool[,1]=e_dirichlet[,1]
        w_diri_pool[,1]=w_dirichlet[,1]
        if (nAllel[i]>1){
              for(j in 2:nAllel[i]){
                      e_diri_pool[i,j]=e_diri_pool[i,j-1]+e_dirichlet[i,j]
                      w_diri_pool[i,j]=w_diri_pool[i,j-1]+w_dirichlet[i,j]
                      }
              }
}
```

```
E_m_locinumber=matrix(0,100,11)
E_f_locinumber=matrix(0,100,11)
W_m_locinumber=matrix(0,100,11)
W_f_locinumber=matrix(0,100,11)

for (i in 1:Sample_num){
    for (j in 1:11){
        m_rand=runif(1)
        m_k=1
        while ( m_rand>w_diri_pool[j,m_k]){
                m_k=m_k+1
                }
        W_m_locinumber[i,j]=m_k
        f_rand=runif(1)
        f_k=1
        while ( f_rand>w_diri_pool[j,f_k]){
                f_k=f_k+1
                }
        W_f_locinumber[i,j]=f_k
    }
    for (j in 1:11){
             m_rand=runif(1)
             m_k=1
             while ( m_rand>e_diri_pool[j,m_k]){
                    m_k=m_k+1
                    }
             E_m_locinumber[i,j]=m_k
             f_rand=runif(1)
             f_k=1
             while ( f_rand>e_diri_pool[j,f_k]){
                    f_k=f_k+1
                    }
             E_f_locinumber[i,j]=f_k
    }
}

W_freq=matrix(0,11,14)
E_freq=matrix(0,11,14)
W_m_freq=matrix(0,11,14)
E_m_freq=matrix(0,11,14)
W_f_freq=matrix(0,11,14)
E_f_freq=matrix(0,11,14)
for (i in 1:11){
    for(j in 1:nAllel[i]){
```

```
        W_m_temp_freq=0
        E_m_temp_freq=0
        W_f_temp_freq=0
        E_f_temp_freq=0
        for (k in 1:Sample_num){
            if(W_m_locinumber[k,i]==j) W_m_temp_freq=W_m_temp_freq+1
            if(E_m_locinumber[k,i]==j) E_m_temp_freq=E_m_temp_freq+1
            if(W_f_locinumber[k,i]==j) W_f_temp_freq=W_f_temp_freq+1
            if(E_f_locinumber[k,i]==j) E_f_temp_freq=E_f_temp_freq+1
        }
        W_m_freq[i,j]=W_m_temp_freq
        E_m_freq[i,j]=E_m_temp_freq
        W_f_freq[i,j]=W_f_temp_freq
        E_f_freq[i,j]=E_f_temp_freq
    }
}
W_freq=(W_m_freq+W_f_freq)/(2*Sample_num)
E_freq=(E_m_freq+E_f_freq)/(2*Sample_num)
Z_freq=(W_freq+E_freq)/2

Fst_ij=matrix(0,11,14)
for(i in 1:11){
    for(j in 1:14){
        if (Z_freq[i,j]==0 | Z_freq[i,j]==1) Fst_ij[i,j]==0
        else Fst_ij[i,j]=((W_freq[i,j]-Z_freq[i,j])^2+(E_freq[i,j]-Z_freq[i,j])^2)/(Z_freq[i,j]*(1-Z_freq[i,j]))
    }
}
Fst_i=matrix(0,11,1)
for(i in 1:11){
    Fst_i[i]=sum(Fst_ij[i,])/nAllel[i]
}

FST[x]=sum(Fst_i)/11
}

t<-seq(0.1,10,by=0.1)
FST_smooth=matrix(0,t_range,1)
FST_smooth<-lowess(t,FST)$y
plot(t,FST,type="l")
Bind_Fst=matrix(0.062,t_range,1)
lines(t,FST_smooth)
lines(t,Bind_Fst)
```

#test_1.R

```
NullBand<- matrix(scan("C:/Economics/thesis/m/Whale_2/my_data_01.out", n = 245*1001), 245, 1001, byrow = TRUE)
Data_Aij<- matrix(scan("C:/Economics/thesis/m/Whale_1/my_data_20.out", n = 245*101), 245, 101, byrow = TRUE)
Band<-matrix(0,245,1000)
for(i in 1:1000){Band[,i]<-lowess(NullBand[,1001],NullBand[,i])$y}
Aij<-matrix(0,245,100)
for(i in 1:100){Aij[,i]<-lowess(Data_Aij[,101],Data_Aij[,i])$y}
Meanband<-matrix(0,245,7)
for(j in 1:245){Meanband[j,]=quantile(Band[j,],prob=c(0.005,0.025,0.05,0.5,0.95,0.975,0.995))}
MeanAij<-matrix(0,245,1)
for(j in 1:245){MeanAij[j]=sum(Aij[j,])/100}
plot (NullBand[,1001],Meanband[,4],type="l",lty=2,ylim=c(-1.5,0),xlab="days apart",ylab="effect")
title("simultaneous nullbands")
lines(NullBand[,1001],Meanband[,1],lty=5)
lines(NullBand[,1001],Meanband[,7],lty=5)
lines(NullBand[,1001],Meanband[,2],lty=4)
lines(NullBand[,1001],Meanband[,6],lty=4)
lines(NullBand[,1001],Meanband[,3],lty=3)
lines(NullBand[,1001],Meanband[,5],lty=3)
lines(NullBand[,1001],MeanAij,lty=1)
legend(0,-1.1,c("50%","90%","95%","99%"),lty=c(2,3,4,5))

#plot(Agg_Aij[,1001],Agg_Aij[,1])
```

% Ch3

% Whole Programe in R

```
#Observation=days that the whale is observed
#Day_Median=days that moved to a same median 300
#Total_number=the total number of observations
#t=the range of the shape of gamma function

t=10

Observation<-matrix(scan("../Observation.txt"), nrow=10, ncol=18, byrow=TRUE)

#To move the observation data to a same median 300
Day_Median=matrix(0,10,18)
median_data=c(1:10)
Total_number=0
for (i in 1:10){
      j=1
```

```
        while (Observation[i,j]>0){
                j=j+1
                Total_number=Total_number+1
        }
        median_data[i]=floor((Observation[i,1]+Observation[i,j-1])/2)
        for (k in 1:j-1){
                Day_Median[i,k]=Observation[i,k]+300-median_data[i]
        }
}

#temp_ID=a data frame including three parts first: whale_ID, year and day
Whale_ID=c(1:Total_number)
Year=c(1:Total_number)
day=c(1:Total_number)
nTemp=0
for (i in 1:10){
        j=1
        while (Day_Median[i,j]>0){
                nTemp=nTemp+1
                Year[nTemp]=i
                day[nTemp]=Day_Median[i,j]
                j=j+1
         }
}
temp_ID<-data.frame(Whale_ID,Year,day)

#the following is used to figure out the hypothesis
#number_day=the number of observations in each day
dayTemp=sort(day)
number_day=c(1:Total_number)
i=1
j=2
while (j<Total_number+1){
        x=1
        while (dayTemp[i]==dayTemp[j]){
              x=x+1
              j=j+1
        }
        number_day[i:j]=x
        i=j
        j=j+1
}
temp_Data<-data.frame(dayTemp,number_day)
```

```
#Hypothesis function
#Mu=the median of each pulse
#Delta=the width of each pulse
#Range=the range of the each pulse
#temp_Hyp=the hypothesis for the E and W groups
#plot_E_hyp=the sum of the hypothesis for the E-groups to plot
#plot_W_hyp=the hypothesis for the W-group to plot
#p_E_hyp=Probability of E-groups
#p_W_hyp=Probability of W-groups

Mu=c(293,300,306)
Delta=c(3.5,3,5)
Range=c(0.55,0.6,1)

Day_year<-c(1:365)
temp_Hyp<-matrix(0,365,3)

for (i in 270:330){
temp_Hyp[i,1]=Range[1]*dnorm(i-Mu[1],sd=Delta[1])
temp_Hyp[i,3]=Range[3]*dnorm(i-Mu[3],sd=Delta[3])
temp_Hyp[i,2]=Range[2]*dnorm(i-Mu[2],sd=Delta[2])
}
k=5/temp_Hyp[Mu[3],3]
plot_E_hyp<-k*(temp_Hyp[,1]+temp_Hyp[,3])
plot_W_hyp<-k*(temp_Hyp[,2])
p_E_hyp=plot_E_hyp/(plot_E_hyp+plot_W_hyp)
p_W_hyp=plot_W_hyp/(plot_E_hyp+plot_W_hyp)

rm(temp_Hyp)

plot(dayTemp,number_day,xlim=c(270,330),ylim=c(0,5))
lines(Day_year,plot_E_hyp)
lines(Day_year,plot_W_hyp,lty=2)

plot(Day_year,p_E_hyp,xlim=c(270,330),ylim=c(0,1),type="l")
lines(Day_year,p_W_hyp,lty=2)


W_population<-(runif(53)<p_W_hyp[temp_ID$day])

Database<-data.frame(temp_ID,W_population)

#Loci=loci and allele data
#nAllel=number of allels present in a locus
```

```
#Gen_frequency=allel frequencies within loci
#e_gamma=the gamma random of e-group
#w_gamma=the gamma random of w-group
#e_dirichlet=the dirichlet distribution of e-group
#w_dirichlet=the dirichlet distribution of w-group

#Input the Gen data
Loci<-matrix(scan("../Loci.txt"), nrow=22, ncol=14, byrow=TRUE)

#Calculate the Gen_frequency
temp_frequency=matrix(0,11,14)
Gen_frequency=matrix(0,11,14)
e_gamma=matrix(0,11,14)
w_gamma=matrix(0,11,14)
e_dirichlet=matrix(0,11,14)
w_dirichlet=matrix(0,11,14)
e_diri_pool=matrix(0,11,14)
w_diri_pool=matrix(0,11,14)

nAllel=c(1:11)
for (i in 1:11){
    temp_frequency[i,]<-Loci[2*i,]
    nAllel[i]=0
    for (j in 1:14){
        if (temp_frequency[i,j]>0){
            nAllel[i]=nAllel[i]+1
        }
    }
}


sim.time=100
Aij.sim=matrix(0,246,sim.time)
Band.sim=matrix(0,246,sim.time)

for (time in 1:sim.time){

for (i in 1:11){
    for (j in 1:nAllel[i]){
        Gen_frequency[i,j]=temp_frequency[i,j]/sum(temp_frequency[i,])
        e_gamma[i,j]=rgamma(1,shape=t*Gen_frequency[i,j])
        w_gamma[i,j]=rgamma(1,shape=t*Gen_frequency[i,j])
    }
    for (j in 1:nAllel[i]){
```

```
           e_dirichlet[i,j]=e_gamma[i,j]/sum(e_gamma[i,])
           w_dirichlet[i,j]=w_gamma[i,j]/sum(w_gamma[i,])
           }
     e_diri_pool[,1]=e_dirichlet[,1]
     w_diri_pool[,1]=w_dirichlet[,1]
     if (nAllel[i]>1){
         for(j in 2:nAllel[i]){
               e_diri_pool[i,j]=e_diri_pool[i,j-1]+e_dirichlet[i,j]
               w_diri_pool[i,j]=w_diri_pool[i,j-1]+w_dirichlet[i,j]
               }
         }
}


#To determine the loci number
#m_locinumber=the locinumber of each whale coming from mother
#f_locinumber=the locinumber of each whale coming from father
m_locinumber=matrix(0,Total_number,11)
f_locinumber=matrix(0,Total_number,11)
for (i in 1:Total_number){
    if (W_population[i]==T){
         for (j in 1:11){
             m_rand=runif(1)
             m_k=1
             while ( m_rand>w_diri_pool[j,m_k]){
                    m_k=m_k+1
                    }
             m_locinumber[i,j]=m_k
             f_rand=runif(1)
             f_k=1
             while ( f_rand>w_diri_pool[j,f_k]){
                    f_k=f_k+1
                    }
             f_locinumber[i,j]=f_k
             }
        }
     else{
         for (j in 1:11){
             m_rand=runif(1)
             m_k=1
             while ( m_rand>e_diri_pool[j,m_k]){
                    m_k=m_k+1
                    }
             m_locinumber[i,j]=m_k
```

```
          f_rand=runif(1)
          f_k=1
          while ( f_rand>e_diri_pool[j,f_k]){
               f_k=f_k+1
               }
          f_locinumber[i,j]=f_k
          }
      }
}

locinumber<-data.frame(m_locinumber,f_locinumber)
Data_loci<-data.frame(temp_ID,W_population,locinumber)

#calculate expected homozygosity at each locus
allele.frequency.sim<-as.list(1:11)
for (i in 1:11){
    allele.frequency.sim[[i]]<-table(c(m_locinumber[,i],f_locinumber[,i]))
    allele.frequency.sim[[i]]<-allele.frequency.sim[[i]]/sum(allele.frequency.sim[[i]])
    }

h.sim=matrix(0,1,11)
for (i in 1:11){
    h.sim[i]=sum(allele.frequency.sim[[i]]^2)
    }

dayapart<-matrix(0,246,1)
temp_matrix<-matrix(0,2,2)
temp_X<-matrix(0,1,11)
temp_Aij<-matrix(0,1,11)
num=0
m=1
for (i in 1:10){
    while(Year[m]==i & m<Total_number){
       n=m+1
       while(Year[n]==i & n<Total_number+1){
           for (k in 1:11){
               temp_matrix[1,]<-c(m_locinumber[m,k],f_locinumber[m,k])
               temp_matrix[2,]<-c(m_locinumber[n,k],f_locinumber[n,k])
               temp_X[k]=0
                       for (col_1 in 1:2){
                       for (col_2 in 1:2){
                           if (temp_matrix[1,col_1]==temp_matrix[2,col_2]){
                           temp_X[k]=temp_X[k]+1
                            }
```

```
                                }
                                }
                        if(h.sim[k]==1){temp_Aij[k]=0}
                          else{temp_Aij[k]=(h.sim[k]-temp_X[k]/4)/(1-h.sim[k])}
                    }
                    num=num+1
                    dayapart[num,1]=day[n]-day[m]
                    Aij.sim[num,time]=sum(temp_Aij)/11
                    n=n+1
                }
        m=m+1
      }
}


#The H0 hypothesis, there is only group e and give the band
#m_locinumber=the locinumber of each whale coming from mother
#f_locinumber=the locinumber of each whale coming from father

H0_m_locinumber=matrix(0,Total_number,11)
H0_f_locinumber=matrix(0,Total_number,11)
for (i in 1:Total_number){
              for (j in 1:11){
                 m_rand=runif(1)
                 m_k=1
                 while ( m_rand>e_diri_pool[j,m_k]){
                          m_k=m_k+1
                          }
                 H0_m_locinumber[i,j]=m_k
                 f_rand=runif(1)
                 f_k=1
                 while ( f_rand>e_diri_pool[j,f_k]){
                          f_k=f_k+1
                          }
                 H0_f_locinumber[i,j]=f_k
                 }

}

H0_locinumber<-data.frame(H0_m_locinumber,H0_f_locinumber)
H0_Data_loci<-data.frame(temp_ID,W_population,H0_locinumber)

#calculate expected homozygosity at each locus
H0_allele.frequency.sim<-as.list(1:11)
```

```
for (i in 1:11){
        H0_allele.frequency.sim[[i]]<-table(c(H0_m_locinumber[,i],H0_f_locinumber[,i]))
        H0_allele.frequency.sim[[i]]<-H0_allele.frequency.sim[[i]]/sum(H0_allele.frequency.sim[[i]])
        }

H0_h.sim=matrix(0,1,11)
for (i in 1:11){
        H0_h.sim[i]=sum(H0_allele.frequency.sim[[i]]^2)
        }


temp_matrix<-matrix(0,2,2)
temp_X<-matrix(0,1,11)
temp_Aij<-matrix(0,1,11)
num=0
m=1
for (i in 1:10){
        while(Year[m]==i & m<Total_number){
              n=m+1
              while(Year[n]==i & n<Total_number+1){
                      for (k in 1:11){
                              temp_matrix[1,]<-c(H0_m_locinumber[m,k],H0_f_locinumber[m,k])
                              temp_matrix[2,]<-c(H0_m_locinumber[n,k],H0_f_locinumber[n,k])
                              temp_X[k]=0
                                              for (col_1 in 1:2){
                                              for (col_2 in 1:2){
                                                      if (temp_matrix[1,col_1]==temp_matrix[2,col_2]){
                                                      temp_X[k]=temp_X[k]+1
                                                        }
                                              }
                                              }
                              if(H0_h.sim[k]==1){temp_Aij[k]=0}
                                else{temp_Aij[k]=(H0_h.sim[k]-temp_X[k]/4)/(1-H0_h.sim[k])}
                      }
                      num=num+1
                      Band.sim[num,time]=sum(temp_Aij)/11
                      n=n+1
                }
        m=m+1
      }
}
}
```

```
#Analyse the data
z<-order(dayapart)
Data.sim<-matrix(0,245,sim.time)
Day.sim<-matrix(0,245,1)
H0.Data.sim<-matrix(0,245,sim.time)
for(i in 1:245)
    {Day.sim[i]<-dayapart[z[i]]
     Data.sim[i,]=Aij.sim[z[i],]
     H0.Data.sim[i,]=Band.sim[z[i],]
    }

H1.Aij<-matrix(0,245,sim.time)
for(i in 1:sim.time){H1.Aij[,i]<-lowess(Day.sim,Data.sim[,i])$y}

H0.Aij<-matrix(0,245,sim.time)
for(i in 1:sim.time){H0.Aij[,i]<-lowess(Day.sim,H0.Data.sim[,i])$y}

Meanband<-matrix(0,245,7)
for(j in 1:245){Meanband[j,]=quantile(H0.Aij[j,],prob=c(0.005,0.025,0.05,0.5,0.95,0.975,0.995))}
MeanAij<-matrix(0,245,1)
for(j in 1:245){MeanAij[j]=sum(H1.Aij[j,])/sim.time}
plot (Day.sim,Meanband[,4],type="l",lty=2,ylim=c(-0.3,0.3),xlab="days apart",ylab="effect")
title("simultaneous nullbands")
lines(Day.sim,Meanband[,1],lty=5)
lines(Day.sim,Meanband[,7],lty=5)
lines(Day.sim,Meanband[,2],lty=4)
lines(Day.sim,Meanband[,6],lty=4)
lines(Day.sim,Meanband[,3],lty=3)
lines(Day.sim,Meanband[,5],lty=3)
lines(Day.sim,MeanAij,lty=1)
legend(0,-1.1,c("50%","90%","95%","99%"),lty=c(2,3,4,5))

plot(Day.sim,MeanAij,type="l",col=2,ylim=c(-0.22,0.12),xlab="days apart",ylab="effect")
for(j in 1:20) lines(Day.sim,H1.Aij[,j],lty=3)

v1<-apply(H1.Aij,2,var)
v2<-apply(H0.Aij,2,var)
v1<-sort(v1)
v2<-sort(v2)
power<-rank(c(v1,v2))[1:sim.time]/(2*sim.time)
plot(ppoints(sim.time),power)
```

```
% Ch5
% Fish Management: Pella_Tomlinson_seq.m
clear;
year=100;
x1(1)=8000;
x2(1)=8000;
r_num=0.03;
k=50000;
z1=1;
z2=5;
for i=1:year-1
    r(i)=r_num+randn*0.2;
    x1(i+1)=x1(i)+r(i)*x1(i)*(1-(x1(i)/k)^z1);
    x2(i+1)=x2(i)+r(i)*x2(i)*(1-(x2(i)/k)^z2);
end
j=1:year;
figure(1)
plot(j,x1,'-',j,x2,'--');
%%
```

```
%Ch5
%Fish Management: constant and variable.m
clear;
r_num=0.03;
year=200;
K=20000;
Y(1)=8000;
x(1)=Y(1);
C_num=144;
C(1)=C_num;
for t=2:year
    r(t)=r_num;%+randn*0.2
    Y(t)=x(t-1)+r(t)*x(t-1)*(1-x(t-1)/K);
    x(t)=Y(t)-C(t-1);
    C(t)=C_num;
end
j=1:year;
figure(1);
plot(j,Y,'-',j,C,'--');
axis([0,200,0,20000]);
%%
```

```
%Ch5
% Fish Management: statistics.m
clear;
r_num=0.03;
year=200;
K=20000;
Y(1)=8000;
x(1)=Y(1);
for C_num=1:140
    lesstime(C_num)=0;
    t=2;
    for cycletime=1:50
        while ((t<=year)&(Y(t-1)>0))
            r(t)=r_num+randn*0.1;
            Y(t)=x(t-1)+r(t)*x(t-1)*(1-x(t-1)/K);
            x(t)=Y(t)-C_num;
            t=t+1;
        end
        if Y(t-1)<0
            lesstime(C_num)=lesstime(C_num)+1;
        end
        t=2;
    end
    ratio(C_num)=lesstime(C_num)/cycletime;
end
figure
j=1:140;
plot(j,ratio,'*')
```

Reference

1. Anonymous. 2005. Workshop on Bowhead Whale Stock Structure Studies in the Beering-Chukchi-Beaufort Seas: 2005-2006. Workshop report, NOAA/AFSC/National Marine Mammal Laboratory and NSB/Department of Wildlife Management, 11 April 2005.
2. Anonymous. 2006. Engineering statistics handbook, National Institute of Standards and Technology (NIST)
3. Bockstoce J (1986) *Whales, ice, and men: the history of whaling in the western Arctic*. University of Washington Press, Seattle, WA.
4. Bockstoce JR, Burns JJ (1993) Commercial whaling in the north pacific sector. In: *The Bowhead Whale* (Eds. Burns JJ, Montague JJ, Cowles CJ), pp.563-577. Society for Marine Mammalogy, Lawrence, KS.
5. Bogoslovskaya, L. S. 2003. The bowhead whale off Chukotka: Integration of scientific and traditional knowledge. Pp. 209-54. In: A. P. McCartney (ed.) Indigenous Ways to the Present: Native Whaling in the Western Arctic. The University of Utah Press, Salt Lake City, UT
6. C.D. Kolstad, *Environmental Economics.* Oxford University Press, 2000.
7. Hardy, G. H. 1908. *Mendelian proportions in a mixed population.* Science 28: 49-50
8. Hotelling, H: *The economics of exhaustible resources*, Journal of Political Economy 39 (2).
9. Jeffrey K. Conner, Michigan State University, and Daniel L. Hartl*, Harvard University, February 2004, A Primer of Ecological Genetics*.

10. Jorde, P. E., T. Schweder, J. W. Bickham, G. H. Givens, G. Suydam, and N.C.Stenseth. 2006. Detecting genetic structure in migrating bowhead whales off the coast of Barrow, Alaska.

11. Jorde, P. E., T. Schweder, and Stenseth N. C. 2004, The Beering-Chukchi-Beaufort stock of bowhead whales: one homogeneous population? IWC/SC/56/BRG36

12. Kevin J. Gaston & John I. Spicer. 2004. "Biodiversity: an introduction", Blackwell Publishing. 2nd Ed., ISBN 1-4051-1857-1

13. LeDuc RG, Dizon Ae, Burdin AM, Blokhin SA, George JC and Brownell RL Jr (2005) Genetic analyses (mtDNA and microsatellites) of Okhotsk and Bering/Chukchi/Beaufort Seas populations of bowhead whales. J. Cetacean Res. Manage. 7(2): 107-111.

14. Melnikov, V. V., Litovka, D. I., Zagrebin, I. A., Zelensky, G.M. and Ainana, L.I. 2004. Shore-based counts of bowhead whales along the Chukotka Peninsula in May and June 1999-2001. Arctic 57(3):290-98.

15. Moore SE, Reeves RR (1993) Distribution and movement. In: The Bowhead Whale Eds. Burns JJ, Montague JJ, Cowles CJ), pp.313-386. Society for Marine Mammalogy, Lawrence, KS.

16. O'Connell, J. (1982) *Welfare Economic Theory*, Auburn House Publishing, Boston

17. Rousset, F. *2000, Genetic differentiation between individuals. J. Evol. Biol. ,13: 58-62.*

18. R. Perman et al, *Natural Resource and Environmental Economics,* Pearson Addison Wesley, 2003.

19. Stern, C.1943. The Hardy-Weinberg law. Science 97:137-138.

20. Varian, Hal R. *Microeconomic Analysis*. W. W. Norton & Company, 3rd Edition