On simulation and Analysis of Stochastic Differential Population Model

Efendi[1]

[1] Mathematics Department, Andalas University, Padang, Indonesia
[1)]e-mail: efendi@fmipa.unand.ac.id

Abstract. In this paper it is studied the population modeling using stochastic differential equation. Population can be predicted with maximum population level probability function. As the function is continues the maximum is stationer type. In validating the stochastic differential population model, it is simulated the Indonesian population census data.The result is that the model well projected Indonesian population in compared to Indonesian Family Planning and Citizen Board (Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN)).

Keywords: Population modeling, Stochastic differential equation

1 Background

It is well known that many ways to model population growth. We can derive model using matrices, differential equation, integral equation, statistics or stochastic differential equation. The advantage on using stochacticity in modelling is that it considers the variying of input or output of model compared to deterministic one. When deterministic is rigid model, stochastic plays flexible manner.

According to Nurdin [3], population growth will depend on four variables that are *natallity*, *mortality*, *imigration* and *emigration*. In modelling population we used to neglect some variable and consider few variables in order to simplify the model. For instance, if we consider closed population, we can neglect migration.

On the other hand, in demoghaphy views, it is known that the growth is also influenced by age distribution, genetic composition and distribution pattern [3]. It is due to population structure also influences the population growth. To simplify the model we can assume that population structure is only determined by population growth rate.

2 Population Modeling

Formula derivation and symbols using in this section refer to Widiowati [5] and Efendi [1]. To understand the modelling process, we rewrite it again here.

Asumption:

1. Population is closed. No emigration or imigration
2. Birth process is random in time interval Δt.
3. Birth process is determined population growth rate λ
4. Population can not be decreased. No death process.
5. The only parameter to consider is population growth rate λ
6. Probability for an emergence of one birth process is proportional to Δt. Let it be $\lambda\Delta t$ which λ is a positive number and probability for no birth in Δt is $1-\lambda\Delta t$.
7. It is not possible for double birth to be happen in Δt.

Stochastic Population Model:

Consider that in beginning there is N_0 population, then birth rate will be $\frac{\Delta N}{N_0 \Delta t}$. It is equal to birth rate λ. Let $P_N(t)$ be probability for population in level N at time t. The probability at time $t+\Delta t$ that population size is still N is $P_N(t+\Delta t)$, How does it happen? There are two possibilities, one is where at beginning there are N population and there is one birth at time interval Δt, and the second, where there are N population but no birth at time interval Δt. Mathematically we can write:

$$P_N(t+\Delta t) = \sigma_{N-1} P_{N-1}(t) + v_N P_N(t) \qquad (1)$$

where σ_{N-1} is probability of one birth among $N-1$ population and v_N is probability for no birth among N population.

If probability of no birth process is $1-\lambda\Delta t$, then probability of no birth among N population is $(1-\lambda\Delta t)^N$, so that $v_N = (1-\lambda\Delta t)^N$. On the other hand, the probability of at least one birth for N population is $1-(1-\lambda\Delta t)^N$. If Δt small enough, then I may assume that there is no double birth among N population. So that $\sigma_{N-1} = 1-(1-\lambda\Delta t)^{N-1}$. Using binomial Newton series, then

$$v_N = (1-\lambda\Delta t)^N = 1 - N\lambda\Delta t + \frac{N(N-1)}{2}(\lambda\Delta t)^2 \mp \cdots$$

$$\sigma_{N-1} = 1-(1-\lambda\Delta t)^{N-1} = 1-(1-(N-1)\lambda\Delta t + \frac{(N-1)(N-2)}{2}(\lambda\Delta t)^2 \mp \cdots$$

Remove non linear parts of $\lambda\Delta t$ then:

$$v_N = 1 - N\lambda\Delta t \qquad (2)$$

$$\sigma_{N-1} = (N-1)\lambda\Delta t \qquad (3)$$

Substitute (2) and (3) to (1), then:

$$P_N(t+\Delta t) = ((N-1)\lambda\Delta t)P_{N-1}(t) + (1-N\lambda\Delta t)P_N(t) \qquad (4)$$

Equation (4) can be rewritten as:

$$\frac{P_N(t+\Delta t)-P_N(t)}{\Delta t}=(N-1)\lambda)P_{N-1}(t)-N\lambda P_N(t) \tag{5}$$

From equation (5), for $\Delta t\rightarrow 0$, then:

$$\lim_{\Delta t\rightarrow 0}\frac{P_N(t+\Delta t)-P_N(t)}{\Delta t}=((N-1)\lambda)P_{N-1}(t)-N\lambda P_N(t)$$

That can be written as:

$$\frac{dP_N(t)}{dt}=\lambda(N-1)P_{N-1}(t)-\lambda NP_N(t)$$

In this case:

$$\frac{dP_N(t)}{dt}+\lambda NP_N(t)=\lambda(N-1)P_{N-1}(t) \tag{6}$$

Equation (6) is a stochastic differential equation (SDE).

Model Solution:

To solve equation (6), let at beginning there is N_0 population and satisfy $P_{N_0}(0)=1$ and $P_N(0)=0,$ for $N\neq N_0$. Let there is no death process for $t>0$, then $P_{N_0-1}(t)=0,$ for $t>0$. So that equation (6) can be written as:

$$\frac{dP_{N_0}(t)}{dt}+\lambda NP_{N_0}(t)=0 \tag{7}$$

The solution of equation (7) is: $P_{N_0}(t)=P_{N_0}(0)e^{-\lambda N_o t}=e^{-\lambda N_o t}$

With the same argument I can calculate $P_{N_0+1}(t)$. Substitute $N=N_0+1$ into SDE, then:

$$\frac{dP_{N_0+1}(t)}{dt}+\lambda NP_{N_0+1}(t)=\lambda(N_0)P_{N_0}(t)=\lambda(N_0)e^{-\lambda N_o t} \tag{8}$$

Solution of equation (8) is:

$$P_{N_0+1}(t)=(N_0)e^{-\lambda N_o t}(1-e^{-\lambda t})$$

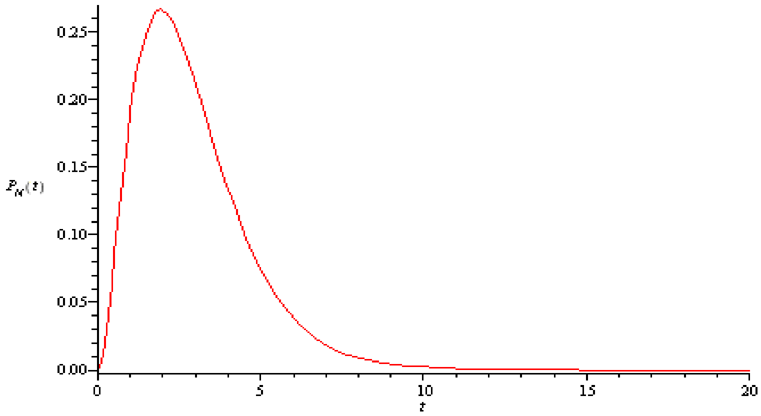
Inductively:

$$P_{N_0+k}(t)=\frac{(N_0)(N_0+1)\ldots(N_0+k-1)}{k!}e^{-\lambda N_o t}(1-e^{-\lambda t})^k \tag{9}$$

Equation (9) can be written as following:

$$P_{N_0+k}(t) = \binom{N_0 + k - 1}{k} e^{-\lambda N_o t}(1 - e^{-\lambda t})^k, \qquad k = 0,1,2,\ldots \qquad (10)$$

To understand the formula in (!0), draw the graphic bellow:



Graph 1. Graphic of probability of population in time *t*.

In Graph.1, the probability for population to be at level N changes follows the time. It means that it is imposible for population to stays at level N every time. It is increased until it reaches maximum and theni t decreased to zero as time tends to infinity. Therefore it is considered that the population will be at size N when the probability is maximums.

Formula for population projection:

Population will be at size $N = N_0 + k$ when $P_{N_0+k}(t)$ reaches maximum, that is when the first derivative of $P_{N_0+k}(t)$ is zero.

$$\frac{dP_{N_0+k}(t)}{dt} = \binom{N_0 + k - 1}{k} [(-\lambda N_o)e^{-\lambda N_o t}(1 - e^{-\lambda t})^k + ke^{-\lambda N_o t}(1 - e^{-\lambda t})^{k-1}(\lambda e^{-\lambda t})] = 0$$

So that respectively:

$$ke^{-\lambda N_o t}(1 - e^{-\lambda t})^{k-1}(\lambda e^{-\lambda t}) = \lambda N_o e^{-\lambda N_o t}(1 - e^{-\lambda t})^k$$
$$\Leftrightarrow \quad ke^{-\lambda N_o t}(\lambda e^{-\lambda t}) = \lambda N_o e^{-\lambda N_o t}(1 - e^{-\lambda t})$$
$$\Leftrightarrow \quad k(e^{-\lambda t}) = N_o(1 - e^{-\lambda t})$$
$$\Leftrightarrow \quad (N_o + k)e^{-\lambda t} = N_0$$
$$\Leftrightarrow \quad e^{\lambda t} = \frac{(N_o + k)}{N_0}$$

$$\Leftrightarrow \quad t = \frac{1}{\lambda}\ln\left(\frac{N_o + k}{N_0}\right)$$

Therefore, the time needed for population at size N_0 to increase to level $N_0 + k$ is:

$$t^* = \frac{1}{\lambda}\ln\left(\frac{N_o + k}{N_0}\right) \tag{11}$$

In the case of population projection, with formula (11) we can also determine the nearest time to data census by variiying growth rate parameter λ. From the formula we know that time is ineverse proportional to λ. Therefore, in simulation, if we start with small λ and we get the time is far enough from time census, then increase it to get closer to time census. With the same argument, if we start with big λ and get the time is far enough from time census, then decrease it to get closer to time census.

3 Validating Model using Indonesia Census Data

To validate the model derived in section 2, consider Indonesia census data since 1930 to 2010 below:

Table 1. Indonesian data census 1930-2010 (in million people) (source: Badan Sensus Penduduk)

Year	1930	1940	1950	1961	1971	1980	1990	2000	2010
Population level	60.7	NA	NA	97.1	119.2	146.9	178.6	205.1	237.6

The simulation on census data of Indonesian population is using *Maple software*. The result of simulation also can be found in Efendi [1]. In order to do validation, we rewrite it here.

Algorithm to do simulation is:

1. Let N_0 to be the population size of data census 1930 and 1980
2. Determine k (the number of additional population from data census)
3. Vary birth rate in range 1.4 % - 1.8 % (Indonesian birth rate in average).
4. Use equation (11) to determine t^* (time needed to reach a population level)
5. Use Table 1 to validate by comparing t^* with census year.

Using above algorithm we write Mapple command program for simulation as following:

```
restart :

N[0] := 146.9 :
lambda := 0.014;
k := 295.9 − 146.9 :
P := binomial(N[0] + k−1, k) * exp( −lambda
     * N[0] * t) * (1−exp( −lambda * t) ) ^k :
Peluang := evalf(P) :
Tahunawal := 1980;
dalamwaktu := evalf( 1/lambda · ln( (N[0] + k)/N[0] ) );
jumlahpopulasimenjadi := N_0 + k;
Tahunakhir := Tahunawal + dalamwaktu;

plot(P, t = 0 .. 31) :
```

Using above program for census data, we have following table:

Table 2. Simulation result with N_0 = 60.7 million people (census 1930)
(source: Efendi [1])

	N (population level for some lambda)				
Period	Lambda= 0.018	Lambda= 0.017	Lambda= 0.0165	Lambda= 0.015	Lambda= 0.014
1930 an	60.7 1930	60.7 1930	60.7 1930	60.7 1930	60.7 1930
1940 an	70.7 1938.4	70.7 1938.9	70.7 1939	70.7 1940	70.7 1940.9
1950 An	82.7 1947	82.7 1948	82.7 1948.7	82.7 1950	82.7 1952
1961 an	97.1 1956	97.1 1957.6	97.1 1958.4	97.1 1961.3	97.1 1963.5
1971 an	119.2 1967.5	119.2 1969.7	119.2 1970	119.2 1975	119.2 1978
1980 an	146.9 1979	146.9 1981	146.9 1983.	146.9 1988.	146.9 1993.
1990 an	178.6 1990	178.6 1993	178.6 1995	178.6 2001.9	178.6 1995
2000 an	205.1 1997.4	205.1 2001.62	205.1 2003.8	205.1 2011	205.1 2017
2010 an	237.6 2005.8	237.6 2010	237.6 2012	237.6 2020	237.6 2027
2020	307	280.5	268	234.4	X

an	2020	2020	2020	2020	
2030 an	367.3 2030	332.3 2030	316.1 2030	X	X

Information for 2:

	Cannot be validated
	Validated with the nearest census year
	Indonesian population prediction on 2020-2030
X	Not possible
N t^*	Population size Time needed to reach population level at N

Discussion:

From table 2, it is seen that there is continues change of birth rate. There is an increasing birth rate from 1930 to 1990, but decreased from 2000 and 2010. Indonesian population prediction for 2030 is between 316.1 to 367.3 million people, it will depend on the real birth rate at that period. This result is similar to BKKBN prediction publish on Thursday, 1 October 2015 [1,4].

According to Indonesian Family Planning and Citizen Board (Badan Kependudukan dan Keluarga Berencana Nasional (BKKBN)) in *Republika online* newspaper on Thursday, 3 October 2015 [4], the size of Indonesian population reaches 345 million people or more on 2030. It is due to rapid growth of Indonesian population. Based on a survey done by Indonesian Survey Board (Badan Survei Demografi dan Kesehatan Indonesia (SDKI)), the growth of Indonesian population from 1980 to 2000 increases significantly. It is known from Indonesian census from 2000 to 2010, the growth rate increases from 1.45% to 1.49%. [1]

4 Conclusion

From simulation in part 3, we know the population model using stochastic differential equation project Indonesian population growth finely in compared to BKKBN prediction.

References:

1. Efendi, Model Persamaan Diferensial Stokastik untuk Memproyeksikan Pertumbuhan Penduduk Indonesia, Seminar Indo-MS, UNP Padang/
2. Fedotov S. etc. *A stochastic model for population migration and the growth of human settlements during neolithic*, Manchester UK, 2008
3. Nurdin, M. *Ekologi Populasi*. Padang: Universitas Andalas Press. (2003)
4. Republika.co.id, diakses tanggal 1 Oktober 2015 pukul 15.00
5. Widiowati and Sutimin, *Buku Ajar Pemodelan Matematika*, Jurusan Matematika FMIPA UNDIP, 2007