

THE PROJECTION OF INDONESIA POPULATION BASED ON STOCHASTIC DIFFERENTIAL EQUATION (SDE)

Efendi¹

¹)Mathematics Department of Andalas University, Padang, Indonesia
E-mail: efendi@sci.unand.ac.id

ABSTRACT

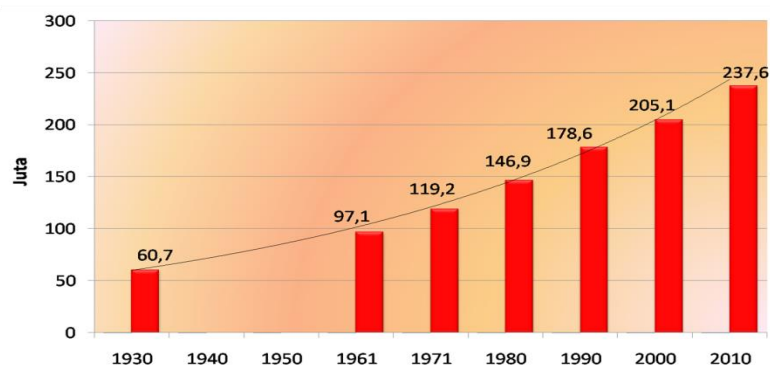
This paper was studied the application of stochastic differential equations to predict the Indonesian population. The data is the Indonesian census from 1930 to 2010 published by The *Indonesian Census Board*, and were simulated using *Maple software*. The result is the Indonesian population growth depends on population growth rate. In the simulation, the growth rate was assumed to vary from 1.4 % to 1.8 %. The result is similar to BKKBN prediction published on Thursday 3 October 2015, that people reached 345 million or more.

Keywords: Stochastic differential equation, the projection of the Indonesian population

INTRODUCTION

According to Indonesian Family Planning and Citizen Board (BKKBN)) in [1] on Thursday, 3 October 2015, the size of the Indonesian population will reach 345 million people or more on 2030. It is due to the rapid growth of the Indonesian growth rate. Based on a survey done by the Indonesian Survey Board (SDKI)), the growth of the Indonesian population from 1980 to 2000 increases significantly. It is known from the Indonesian census from 2000 to 2010, the growth rate increases from 1.45% to 1.49%. [1] In the Demography theory, it is known that population prediction depends on population age structure. In this case, from census data in 1961, it is known that the young population is dominant, which is quite different from the data census in 2010, where the old population is dominant. So that growth projection may change. In 1961, the amount of the Indonesian population is 97.1 million people. Meanwhile, in 2010, a great amount of the population reaches 237.6 million people. It means that every 10 years, there is an additional population of around 32 million people or 2.6 million people yearly. So that, the prediction of the Indonesian population is 245.4 million, in 2013 and is 248 million people in 2014 [1].

The results of the Indonesian census from 1930 to 2010 is shown in graphic 1.



Graphic 1. The amount of Indonesian population in each census period

Generally, the population growth usually will depend on four variables that are *natality*, *mortality*, *immigration*, and *emigration*. Besides that, the growth is also influenced by the age distribution, the genetic composition, and the distribution pattern [2]. But generally, the information of the four variables and population structure is included in the population growth rate. Therefore in this paper, I will only use the growth rate to predict an Indonesian population growth. The growth rate used here varies from 1.4 % to 1.8 % based on the average growth rate of the Indonesian population [1].

THEORITICAL FRAMEWORK

The Stochastic differential population model

The formula derivation and symbols using in this section refer to [1]. Let the birth process is random in the time interval Δt . The probability of the emergence of one birth process is proportional to Δt . Let it be $\lambda \Delta t$ which λ is a positive number. Let it is not possible for double birth to be happen in Δt . So that probability for no birth in Δt is $1 - \lambda \Delta t$. If at 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 . So that λ is not only a probability of one birth but also it is a birth rate.

Let $P_N(t)$ be a probability for population in size N at time t . I can calculate probability at a time $t + \Delta t$ that population size is still N . To calculate $P_N(t + \Delta t)$, 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 I can write the two possibilities as:

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

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 the probability of no birth process is $1 - \lambda \Delta t$, then the probability of no birth among N population is $(1 - \lambda \Delta t)^N$, so that I have $v_N = (1 - \lambda \Delta t)^N$. 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, I have:

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

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

If I neglect the nonlinear parts of $\lambda \Delta t$, then I have:

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

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

Substitute equation (2) and (3) to (1), I have:

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

Equation (4) can be rewritten as:

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

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

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

That can be written as:

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

In this case I have:

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

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

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:

$$\frac{dP_{N_0}(t)}{dt} + \lambda N P_{N_0}(t) = 0 \quad (7)$$

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

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

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

The solution of equation (8) is:

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

Inductively, I can have:

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

Equation (9) can be written as following:

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

The population will be at size $N = N_0 + k$ when $P_{N_0+k}(t)$ getting maximum, which is at the first derivative of $P_{N_0+k}(t)$ is zero.

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

So that respectively I have:

$$\begin{aligned} k e^{-\lambda N_0 t} (1 - e^{-\lambda t})^{k-1} (\lambda e^{-\lambda t}) &= \lambda N_0 e^{-\lambda N_0 t} (1 - e^{-\lambda t})^k \\ \Leftrightarrow k e^{-\lambda N_0 t} (\lambda e^{-\lambda t}) &= \lambda N_0 e^{-\lambda N_0 t} (1 - e^{-\lambda t}) \\ \Leftrightarrow k (e^{-\lambda t}) &= N_0 (1 - e^{-\lambda t}) \end{aligned}$$

$$\begin{aligned} \Leftrightarrow (N_0 + k)e^{-\lambda t} &= N_0 \\ \Leftrightarrow e^{\lambda t} &= \frac{(N_0 + k)}{N_0} \\ \Leftrightarrow t &= \frac{1}{\lambda} \ln \left(\frac{N_0 + k}{N_0} \right) \end{aligned}$$

So the time needed for population at size N_0 to increase at size $N_0 + k$ is:

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

DISCUSSION

The simulation to predict Indonesian population projection

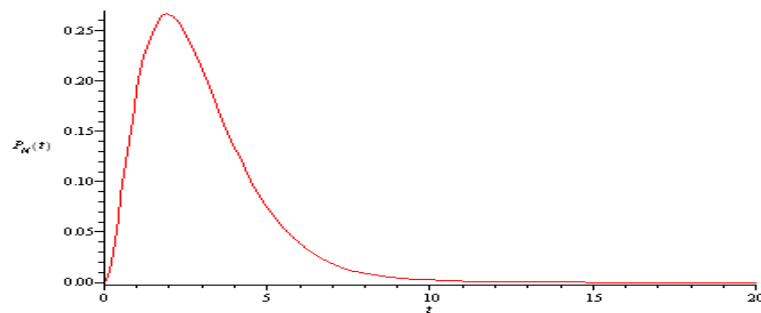
In this section, I simulate census data of the Indonesian population using *Maple* software based on formula derive in section2.

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.

Table 1. Indonesian data census 1930-2010 (in million people)

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



Graph 1. Graphic of probability of population in time.

So that I can write the mapple command program 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 the above program for census data year 1930 as beginning population (60.7 million people), I have following table:

Table 2. Simulation result with $N_0 = 60.7$ million people (census 1930)

	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 an	307 2020	280.5 2020	268 2020	234.4 2020	X
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	Population size
t*	Time needed to reach population level at N

From table 2, it is seen that there is a continues change of birth rate. There is an increasing birth rate from 1930 to 1990, but decreased from 2000 and 2010. 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 published on Thursday, 1 October 2015.

Next with census data 1980 as beginning population (146.9 million people) from simulation I have:

Table 3. Simulation result with $N_0 = 146.9$ million people (census data 1980)

	N (Population size for some lambda) (in million people)				
Period	Lambda = 0.018	Lambda = 0.017	Lambda = 0.0165	Lambda = 0.015	Lambda = 0.014
1980 an	146.9 1979	146.9 1981	146.9 1983.	146.9 1988.	146.9 1993.

1990 an	178.6 1990.8	178.6 1991.8	178.6 1991.8	178.6 2001.9	178.6 1995
2000 an	205.1 1998.5	205.1 1999.6	205.1 2000.2	205.1 2002.2	205.1 2003.8
2010 an	237.6 2006.7	237.6 2008	237.6 2009	237.6 2012	237.6 2014
2020 an	301.8 2020	290 2020	284.3 2020	268 2020	257.5 2020
2030 an	361.4 2030	343.7 2030	335.3 2030	311 2030	295.9 2030

Information for table 3:

	Validated with the nearest census year
	Indonesian population prediction on 2020-2030
N	Population size
t^*	Time needed to reach population level at N

From table 3, the result is similar to table 2, that there is an increase of birth rate from 1930 to 1990, but there is a decrease from 2000 to 2010. The prediction for 2030 is between 295.9 to 361.4 million people, it depends on the real birth rate at that period. This result is similar to BKKBN prediction.

CONCLUSION

From simulation based on SDE, it is known that the Indonesian population prediction depends on birth rate. It is known the birth rate of Indonesian population is varied for every period. Using birth rate from 1.4% to 1.8% the prediction of the Indonesian population is 350 million people at 2030.

REFERENCES

- [1] Fedotov S. etc. *A stochastic model for population migration and the growth of human settlements during neolithic*, Manchester UK, 2008
- [2] Nurdin, M. *Ekologi Populasi*. Padang: Universitas Andalas Press. (2003)
- [3] Republika.co.id, accessed on 1 Oktober 2015 at 15.00
- [4] Widiowati and Sutimin, *Buku Ajar Pemodelan Matematika*, Jurusan Matematika FMIPA UNDIP, 2007