iBusiness, 2022, 14, 31-40
https://www.scirp.org/journal/ib
ISSN Online: 2150-4083
ISSN Print: 2150-4075

Exlog Weighted Sum Model for Long Term Forecasting

Luvsandash Ouyntsetseg[1], Natsagdorj Tungalag[1], Rentsen Enkhbat[2*]

[1]Department of Accounting, School of Business, National University of Mongolia, Ulaanbaatar, Mongolia
[2]Institute of Mathematics and Digital Technology, Mongolian Academy of Sciences, Ulaanbaatar, Mongolia
Email: *renkhbat46@yahoo.com

How to cite this paper: Ouyntsetseg, L., Tungalag, N., & Enkhbat, R. (2022). Exlog Weighted Sum Model for Long Term Forecasting. *iBusiness, 14,* 31-40.
https://doi.org/10.4236/ib.2022.142003

Received: March 16, 2022
Accepted: April 25, 2022
Published: April 28, 2022

Abstract

Long-term forecasting of key macroeconomic indicators such as population is very important for future development policy-making. Population plays an important role in economic decision-making, social security and economic growth. So it is important to develop a good model for predicting economic indicators. In order to improve the growth model, we introduce a new model called Exlog Weighted Sum Model for predicting macroeconomic indicators. This model combines both exponential and logistic models. The proposed model was tested for predicting Mongolian population up to 2040.

Keywords

Exponential Model, Logistic Model, Exlog Weighted Sum Model, Forecasting

1. Introduction

Population growth is described as a function of time usually by dynamic models based on differential equations.

The most common practical methods are component, exponential and logistic models (Enkhbat & Tungalag, 2006).

On the other hand, the population growth can be considered stochastic variable since population depends on social and economic policies, political stability and so on.

In this paper, first we examine the existing growth models such as exponential, logistic and stochastic models. Then we deal with two classic models. It is well known that exponential and logistic functions have some drawbacks in terms of too fast growth or slow growth in a long term prediction in economy. For this purpose, we introduce a new model called Exlog Weighted Sum Model for predicting macroeconomic indicators and focus on the parameter estimation problem of the exponential and logistic models. Finding a parameter of the exponential model reduces to unconstrained minimization problem which has been solved analytically. As far as the parameter of logistic model is concerned, the least square method reduces to difficult nonconvex optimization problem but we have proposed simple formula for finding parameters of a new model. The parameters of both exponential and logistic models have been estimated for Mongolian population data. We also propose a new method for estimating optimal weights of the Exlog model using statistical data. The new model was tested on Mongolian population forecasting.

2. Exponential and Logistic Models

The general form to model the growth of population is

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y),$$

where y is equal to the number of individuals at time t, i.e. the population.

Many different forms of $f(y)$ exist. Two simple models are exponential and logistic models which are the most popular models (Panik, 2014) for predicting economic indicators.

The exponential model is given by the following differential equation:

$$\begin{cases} \frac{\mathrm{d}y}{\mathrm{d}t} = ry, \\ y(0) = y_0. \end{cases}$$

where r-growth rate, t-period, y_0-population for the base year.

This model plays also an important role in Solow growth theory (Solow, 1956) as a function of population growth.

By solving this equation, we get

$$y(t) = y_0 \mathrm{e}^{rt} \tag{1}$$

Using the least square approximation, we estimated the growth rate r based on data in **Table 1**.

For estimating parameter r using statistical data, we apply the least square method reducing this problem to an unconstrained convex minimization problem after taking logarithm from the expression (1).

In order to estimate parameter r of the exponential model

$$y = y_0 \mathrm{e}^{rt}$$

We use the least square method after reducing it to linear model

$$\ln y = \ln y_0 + rt$$

$$F(r) = \sum_{i=1}^{n} \left[\ln y_0 + rt_i - y_i\right]^2 \to \min \tag{2}$$

Denote by $\overline{y}_0 = \ln y_0$ and solve problem (2)

Table 1. Mongolian population data from 2000 to 2020.

Year	Thousand people
2000	2407.5
2001	2442.5
2002	2475.4
2003	2504.0
2004	2533.1
2005	2562.4
2006	2583.2
2007	2620.9
2008	2665.9
2009	2716.3
2010	2761.0
2011	2811.7
2012	2867.7
2013	2930.3
2014	2995.9
2015	3057.8
2016	3119.9
2017	3177.9
2018	3238.5
2019	3296.9
2020	3357.5

Source: http://www.1212.mn/ (accessed on 14 January, 2022).

$$F'(r) = 2\sum_{i=1}^{n}\left(\overline{y}_0 + rt_i - y_i\right)t_i = 0$$

$$\sum_{i=1}^{n}\overline{y}_0 t_i + r\sum_{i=1}^{n}t_i^2 - \sum_{i=1}^{n}y_i t_i = 0$$

From the last equation, we find

$$r = \frac{\sum_{i=1}^{n} y_i t_i}{\sum_{i=1}^{n} t_i^2} \tag{3}$$

For our data r is computed as 0.0167.

On the other hand, the logistic model is given by the following differential equation:

$$y' = ky(M - y)$$
$$y(0) = y_0$$

The solution to this equation is

$$y(t)=\frac{My_0}{y_0+(M-y_0)\mathrm{e}^{-Mkt}}$$

This function called the logistic curve which is more suitable for modeling long-term growth.

The parameters M and k of the model are estimated usually by the least square method. For this purpose, the following type of a statistical data is used.

Year	t_i	t_1	t_2	…	t_m
MLS	y_i	y_1	y_2	…	y_m

$$F(M,k)=\sum_{i=1}^{m}\left[\frac{My_0}{y_0+(M-y_0)\mathrm{e}^{-Mkt_i}}-y_i\right]^2\to\min,\quad M>0,k>0 \tag{4}$$

Problem (4) is a constrained minimization problem has to be solved with respect to variables M and k. From a view point of optimization theory and methods, it is well known that this is very difficult nonconvex optimization problem. $F(M, k)$ is a differentiable nonlinear function. In general, problem (4) can be solved numerically by the gradient method. The gradient of the function is computed as follows.

$$\frac{\partial F}{\partial M}=2\sum_{i=1}^{m}\left[\frac{(p_i-f(M,k,t_i))\cdot\mathrm{e}^{-Mkt_i}\left[1-(M-p_0)kt_i-p_0(M-p_0)-p_0^2\right]}{\left(p_0+(M-p_0)\mathrm{e}^{-Mkt_1}\right)^2}\right]$$

$$\frac{\partial F}{\partial k}=2\sum_{i=1}^{m}\left[p_i-f(M,k,t_i)\right]\left(\frac{\mathrm{e}^{-Mkt_i}\left[1-(M-p_0)(kt_i+p_0)\right]-p_0^2}{\left(p_0+(M-p_0)\mathrm{e}^{-Mkt_i}\right)^2}\right)$$

$$gradF(M,k)=\left(\frac{\partial F}{\partial M},\frac{\partial F}{\partial k}\right)$$

The maximum value of M can be estimated by exponential growth in advance. If the maximum value M is given, problem (4) reduces to one dimensional global minimization problem with respect to k.

$$\varphi(k)=\sum_{i=1}^{m}\left[\frac{My_0}{y_0+(M-y_0)\mathrm{e}^{-Mkt_i}}-y_i\right]^2\to\min,\quad k>0 \tag{5}$$

In order to estimate M and k by the least square method, we propose the following simple method. The first order differential equation of the model should be replaced by the difference scheme.

$$y'\approx y(t+1)-y(t);$$
$$y(0)=y_0$$

The equation has the following form.

$$y(t+1)-y(t)=ky(M-y)$$
$$y(t+1)=Mky-ky^2+y(t)$$

We have to find k and M models, using the least square method. In this case, the following function has to be minimized:

$$F(M,k)=\sum_{i=1}^{m}\left[Mky_i-ky_i^2+y_i-y_{i+1}\right]^2$$

This problem is two variables' minimization problem, and optimal conditions for extremum is written as follows:

$$\frac{\partial F}{\partial M}=2\sum_{i=1}^{m}\left[Mky_i-ky_i^2+y_i-y_{i+1}\right]ky_i=0$$
$$\frac{\partial F}{\partial k}=2\sum_{i=1}^{m}\left[Mky_i-ky_i^2+y_i-y_{i+1}\right]\left(My_i-y_i^2\right)=0$$

If it is simplified:

$$\frac{\partial F}{\partial M}=\sum_{i=1}^{n}\left(Mky_i^2-ky_i^3+y_i^2-y_iy_{i+1}\right)=0;$$
$$\frac{\partial F}{\partial k}=\sum_{i=1}^{m}\left[\left(Mky_i-ky_i^2+y_i-y_{i+1}\right)My_i-\left(Mky_i-ky_i^2+y_i-y_{i+1}\right)y_i^2\right]=0$$

Simplification of the system's second equation is:

$$\sum_{i=1}^{m}\left(M^2ky_i^2-Mky_i^3+My_i^2-My_iy_{i+1}-Mky_i^3+ky_i^4-y_i^3+y_i^2y_{i+1}\right)=0$$

which gives the following equation:

$$M^2k\sum_{i=1}^{m}y_i^2-2Mk\sum_{i=1}^{m}y_i^3+M\sum_{i=1}^{m}\left(y_i^2-y_iy_{i+1}\right)+k\sum_{i=1}^{m}y_i^4=\sum_{i=1}^{m}\left(y_i^3-y_i^2y_{i+1}\right).$$

Then the system becomes as following:

$$Mk\sum_{i=1}^{m}y_i^2-k\sum_{i=1}^{m}y_i^3=\sum_{i=1}^{m}\left(y_iy_{i+1}-y_i^2\right);$$
$$M^2k\sum_{i=1}^{m}y_i^2-2Mk\sum_{i=1}^{m}y_i^3+M\sum_{i=1}^{m}\left(y_i^2-y_iy_{i+1}\right)+k\sum_{i=1}^{m}y_i^4=\sum_{i=1}^{m}\left(y_i^3-y_i^2y_{i+1}\right)$$

This system of nonlinear equations should be solved with respect to M and k. If M is given in advance, the system reduces to one variable equation:

$$\frac{\partial F}{\partial k}=\sum_{i=1}^{m}\left[Mky_i-ky_i^2+y_i-y_{i+1}\right]\left(My_i-y_i^2\right)=0.$$

We find k from the equation:

$$k\left[\bar{M}^2\sum_{i=1}^{m}y_i^2-2\bar{M}\sum_{i=1}^{m}y_i^3+\sum_{i=1}^{m}y_i^4\right]=\sum_{i=1}^{m}\left(y_i^3-y_i^2y_{i+1}\right)-\bar{M}\sum_{i=1}^{m}\left(y_i^2-y_iy_{i+1}\right),$$

$$\bar{k}=\frac{\sum_{i=1}^{m}\left(y_i^3-y_i^2y_{i+1}\right)-\bar{M}\sum_{i=1}^{m}\left(y_i^2-y_iy_{i+1}\right)}{\bar{M}^2\sum_{i=1}^{m}y_i^2-2\bar{M}\sum_{i=1}^{m}y_i^3+\sum_{i=1}^{m}y_i^4}.$$

In practice, the maximum value of M can be estimated by the exponential growth in advance. And then use an approximation formula of y'.

$$y_{i+1}-y_i=ky_i\left(M-y_i\right),y_i=y\left(t_i\right),i=1,2,\cdots,n$$

$$k_i = \frac{y_{i+1} - y_i}{y_i (M - y_i)}, i = 1, 2, \cdots, n$$

$$\bar{k} = \frac{\sum_{i=1}^{n} k_i}{r} \tag{6}$$

3. The Stochastic Population Model

We introduce uncertainty in the population growth. Due to the unpredictable nature of population growth, we add a random element to the exponential growth equation (Enkhbat, Enkhbayar, & Bayanjargal, 2018).

$$\mathrm{d}y(t) = ry(t)\mathrm{d}t + \sigma y(t)\mathrm{d}W$$

where r is the growth rate, σ is number called diffusion coefficient and W is a Wiener process.

Let the number of population for base year be y_0 at time $t = 0$ and $t_i = i\Delta t$, so the numbers of population are to be determined at discrete points $\{t_i\}$.

Then our discrete-time model is

$$y(t_{i+1}) = y(t_i) + r\Delta t y(t_i) + \sigma\sqrt{\Delta t} Z_i y(t_i),$$

where $Z_i (i = 0, 1, 2, \cdots)$ are i.i.d $N(0,1)$.

Based on the Central Limit theorem and $Z_i (i = 0, 1, 2, \cdots)$ are i.i.d $N(0,1)$, we take a limit $\Delta t \to 0$ to get continuous population model. Continuous time expression for the model is

$$y(t) = y_0 \mathrm{e}^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}Z},$$

where $Z \sim N(0,1)$.

A random variable $y(t)$ of the form has a so-called *lognormal distribution*, that is, its log is normally distributed.

So we can describe the evolution of the population over any sequence of time points $0 = t_0 < t_1 < t_2 < \cdots < t_M$ by

$$y(t_{i+1}) = y(t_i) \mathrm{e}^{\left(r - \frac{1}{2}\sigma^2\right)(t_{i+1} - t_i) + \sigma\sqrt{t_{i+1} - t_i} Y_i},$$

for i.i.d $Y_i \sim N(0,1)$.

4. Exlog Weighted Sum Model

Using the exponential and logistic models, we introduce so-called Exlog Weighted Sum Model given by the following formula:

$$y(t) = \lambda_1 y_0 \mathrm{e}^{rt} + \lambda_2 \frac{My_0}{y_0 + (M - y_0)\mathrm{e}^{-Mkt}} \tag{7}$$

where,

$y(t)$: Macroeconomic indicator;

y_0: initial value of y at the moment $t = 0$;

λ_1, λ_2 are parameters of the model,and $\lambda_1 + \lambda_2 = 1$, $\lambda_1 > 0$, $\lambda_2 \geq 0$.

t	t_i	t_1	t_2	…	t_m
y	y_i	y_1	y_2	…	y_m

Using the above data, we can estimate the parameters of the model by solving the following constrained minimization problem:

$$F(\lambda_1,\lambda_2)=\sum_{i=1}^{m}\left[\lambda_1 y_0 \mathrm{e}^{rt}+\lambda_2\frac{My_0}{y_0+(M-y_0)\mathrm{e}^{-Mkt}}-y_i\right]^2\to\min \quad (8)$$

$$\lambda_1+\lambda_2=1,\ \lambda_1>0,\ \lambda_2\geq 0 \quad (9)$$

In general, problem (8)-(9) is convex optimization problem (Enkhbat, 2018). For simplicity, we assume that we have the following models

$$y^{ex}=y_0\mathrm{e}^{rt},\ y^{log}=\frac{My_0}{y_0+(M-y_0)\mathrm{e}^{-Mkt}}$$

and data:

t	y_i^{ex}	y_i^{log}	y_i
t_0	y_0^{ex}	y_0^{log}	y_0
t_1	y_1^{ex}	y_1^{log}	y_1
t_2	y_2^{ex}	y_2^{log}	y_2
…	…	…	…
t_m	y_m^{ex}	y_m^{log}	y_m

Formula (7) has the form

$$y=y^{ex}+y^{log}.$$

Here, values of y_i^{ex} and y_i^{log} have been estimated by exponential and logistic functions.

Then finding parameters of the Exlog Weighted Sum Model reduces to a constrained minimization problem:

$$F(\lambda_1,\lambda_2)=\sum_{i=1}^{m}\left[\lambda_1 y_i^{ex}+\lambda_2 y_i^{log}-y_i\right]^2\to\min \quad (10)$$

$$\lambda_1+\lambda_2=1,\ \lambda_1\geq 0,\ \lambda_2\geq 0 \quad (11)$$

Since the problem is convex, then we apply Lagrange method. The Lagrangean function is

$$F(\lambda_1,\lambda_2,\lambda)=\sum_{i=1}^{m}\left[\lambda_1 y_i^{ex}+\lambda_2 y_i^{log}-y_i\right]^2+\lambda(\lambda_1+\lambda_2-1)$$

where λ is a Lagrange multiplier.

If we take derivatives of Lagrangean function with respect to variables and equal to zero, we obtain the following system of equations:

$$\begin{cases} \dfrac{\partial L}{\partial \lambda_1} = 2\sum_{i=1}^{m}\left(\lambda_1 y_i^{ex} + \lambda_2 y_i^{log} - y_i\right) y_i^{ex} + \lambda = 0 \\ \dfrac{\partial L}{\partial \lambda_2} = 2\sum_{i=1}^{m}\left(\lambda_1 y_i^{ex} + \lambda_2 y_i^{log} - y_i\right) y_i^{log} + \lambda = 0 \\ \dfrac{\partial L}{\partial \lambda} = \lambda_1 + \lambda_2 = 1 \end{cases}$$

We simplify the above system as follows:

$$\begin{cases} 2\lambda_1 \sum_{i=1}^{m}\left(y_i^{ex}\right)^2 + 2\lambda_2 \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right) = 2\sum_{i=1}^{m}\left(y_i y_i^{ex}\right) - \lambda \\ 2\lambda_1 \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right) + 2\lambda_2 \sum_{i=1}^{m}\left(y_i^{log}\right)^2 = 2\sum_{i=1}^{m}\left(y_i y_i^{\log}\right) - \lambda \\ \lambda_1 + \lambda_2 = 1 \end{cases}$$

Consequently we have,

$$\begin{cases} 2\lambda_1 \sum_{i=1}^{m}\left(y_i^{ex}\right)^2 + 2\left(1-\lambda_1\right) \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right) = 2\sum_{i=1}^{m}\left(y_i y_i^{ex}\right) - \lambda \\ 2\lambda_1 \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right) + 2\left(1-\lambda_1\right) \sum_{i=1}^{m}\left(y_i^{log}\right)^2 = 2\sum_{i=1}^{m}\left(y_i y_i^{log}\right) - \lambda \\ \lambda_2 = 1 - \lambda_1 \end{cases}$$

Hence, we obtain the following system of equations:

$$\begin{cases} 2\lambda_1 \left[\sum_{i=1}^{m}\left(y_i^{ex}\right)^2 - \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right)\right] = 2\sum_{i=1}^{m}\left(y_i y_i^{ex} - y_i^{log} y_i^{ex}\right) - \lambda \\ 2\lambda_1 \left[\sum_{i=1}^{m}\left(y_i^{ex} y_i^{log}\right) - \sum_{i=1}^{m}\left(y_i^{log}\right)^2\right] = 2\sum_{i=1}^{m}\left(y_i y_i^{log} - \left(y_i^{log}\right)^2\right) - \lambda \end{cases}$$

Substituting a value of

$$\lambda_2 = 1 - \lambda_1$$

into last equation, we find λ_1, λ_2 and λ.

$$2\lambda_1 \sum_{i=1}^{m}\left(y_i^{ex}\right)^2 - \sum_{i=1}^{m}\left(y_i^{log} y_i^{ex}\right) + \sum_{i=1}^{m}\left(y_i^{log}\right)^2$$
$$= 2\sum_{i=1}^{m}\left(y_i y_i^{ex} - y_i^{log} y_i^{ex} - y_i y_i^{log} + \left(y_i^{log}\right)^2\right)$$

$$\begin{cases} \lambda_1 = \dfrac{\sum_{i=1}^{m}\left[y_i y_i^{ex} - y_i^{\log} y_i^{ex} - y_i y_i^{log} + \left(y_i^{log}\right)^2\right]}{\sum_{i=1}^{m}\left(y_i^{ex} - y_i^{log}\right)^2} \\ \lambda_2 = 1 - \lambda_1 \end{cases} \quad (12)$$

The values of λ_1, and λ_2 were computed by (12) and found as $\lambda_1 = 0.8$, and $\lambda_2 = 0.2$.

Formula (12) can be used in predicting some other macroeconomic indicators.

5. Forecasting Mongolian Population Growth

Traditional exponential model provides too fast growth, on the other hand, the logistic model ensures too slow growth which are not suitable for long term growth of considered indicators. We illustrate Exlog Weighted Sum Model on Mongolian Population data. Data covers period between 2000 and 2020. In model (7) we take values of $M = 5000.0$ and k, r computed by formulas (3), (6) as follows:

$$k = 0.0000079, \quad r = 0.0167.$$

Based on survey for the period 2000-2020 (Tungalag et al., 2015), we predict the Mongolian population by Exlog Weighted Sum Model:

$$y = 1926e^{0.0167t} + \frac{2407500}{2407.5 + 2592.5e^{-0.0395t}}$$

We forecast the Mongolian population up to the 2040 year in **Table 2**.

6. Conclusion

We examine the existing population growth models such as exponential and logistic models and propose a method for finding parameters of the models by

Table 2. Mongolian population up to the 2040 year (thousands of people).

	y^{ex}	y^{log}	y^{exlog}
2021	3424.3	3403.1	3420.1
2022	3482.2	3445.8	3474.9
2023	3541.1	3487.8	3530.5
2024	3601.1	3529.2	3586.7
2025	3662.0	3569.9	3643.6
2026	3723.9	3610.0	3701.1
2027	3786.9	3649.3	3759.4
2028	3851.0	3688.0	3818.4
2029	3916.2	3725.9	3878.1
2030	3982.4	3763.1	3938.5
2031	4049.8	3799.5	3999.7
2032	4118.3	3835.2	4061.7
2033	4188.0	3870.2	4124.4
2034	4258.8	3904.4	4187.9
2035	4330.9	3937.9	4252.3
2036	4404.1	3970.6	4317.4
2037	4478.7	4002.6	4383.4
2038	4554.4	4033.8	4450.3
2039	4631.5	4064.2	4518.0
2040	4709.8	4093.9	4586.7

solving corresponding convex and nonconvex optimization problems.

We propose also so-called Exlog Weighted Sum Model combining well known classical models such as exponential and logistic for predicting macroeconomic indicators of Mongolia. We derive formulas for estimating weights of the proposed model by solving a constrained convex minimization problem. The proposed model was tested for predicting Mongolian population up to 2040. The proposed approach can be applied to forecasting any macroeconomic indicators.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Enkhbat, R. (2018). *Optimization Methods* (pp. 78-88). MUIS Press Publisher.

Enkhbat, R., & Tungalag, N. (2006). Mathematical Modeling of Pension Insurance of Mongolia. In *Transaction Economy Problems in Central Asia*, *Scientific International Conference Proceedings* (pp. 177-185). National University of Mongolia.

Enkhbat, R., Enkhbayar, J., & Bayanjargal, D. (2018). Application of Stochastic Differential Equations in Population Growth in Optimization Applications. In *Optimization Applications in Economics and Finance* (pp. 82-88). LAP Lambert Academic Publishing.

Panik, M. J. (2014). Fundamentals of Population Dynamics. In *Growth Curve Modeling: Theory and Applications* (pp.352-371). John Wiley.

Solow, R. M. (1956). A Contribution to the Theory of Economic Growth. *Quarterly Journal of Economics, 70,* 65-94. https://doi.org/10.2307/1884513

Tungalag N., Oyuntsesteg L., Enkhbat R., & Enkhbayar J. (2015). An Evaluation of Pension Insurance Reform Options for Mongolia. *Proceedings of China International Conference on Insurance and Risk Management*, 551 p.