CHAPTER 3

Research Methodology

3.1 Conceptual Framework

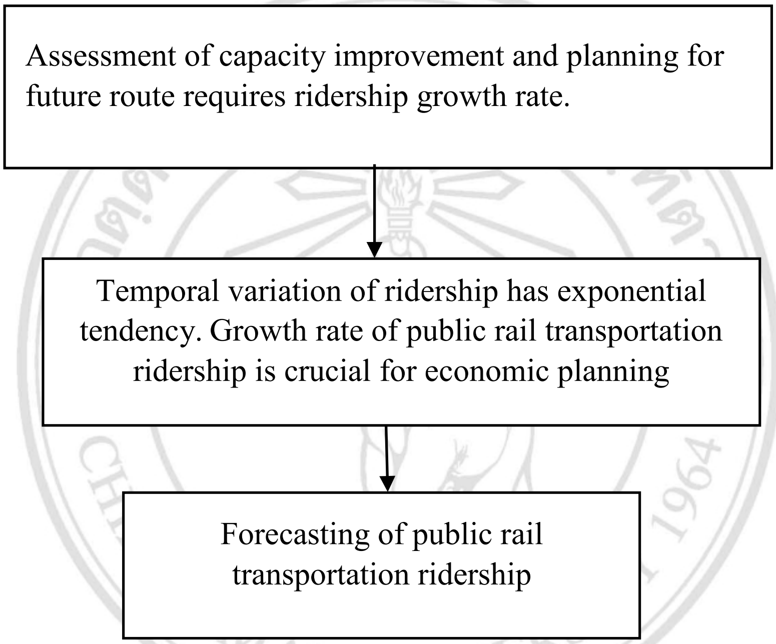

Figure 3.1: Conceptual Framework

3.2 Data collection

To model a public rail transportation use in Bangkok with the stochastic exponential growth equation that is driven by a Standard Brownian Motion. The model identification derived from the Bayesian inference that determines the parameters of the model which are the growth rate (μ) and diffusion term (σ).

Besides, to forecast the ridership growth rate. This study applies the secondary data which is the monthly data of the BTS or sky train ridership from January in 2011 to December in 2016, consisting of 72 observations collected from the Bangkok Mass Transit System Public Company (BTS, 2016).

3.3 Model Specification

To model a public rail transportation use in Bangkok, this study will apply the exponential solution of a stochastic differential equation from a standard Brownian motion and to forecast the growth rate of ridership by using the Markov Chain Monte Carlo method.

For constant drift and diffusion terms, the exponential solution of stochastic differential equation from a standard Brownian motion is

$$X_t = X_0 \, exp\left(\mu t - \frac{\sigma^2}{2} t + \sigma B_t \right) \quad (3.3.1)$$

$$Y_t = \left(\mu t - \frac{\sigma^2}{2} t + \sigma B_t \right) \quad (3.3.2)$$

where

$$Y_t = \ln\left(\frac{S_t}{S_0} \right) \quad (3.3.3)$$

μ and σ are constant

t is time instant

Y_t is the number of ridership at time t

μ is the growth rate of ridership at time t

σ is diffusion term

B_t is one-dimension Brownian motion

The model according to equation (3.2.1)-(3.2.3) will be used in this study.

The drift and diffusion terms in the model will be determined using Bayesian inference.

The posterior distribution equation is

$$f(\mu, \sigma \mid D) = CL(D \mid \mu, \sigma) f(\mu, \sigma) \quad (3.3.4)$$

where

$f(\mu, \sigma \mid D)$ is the posterior distribution

$L(D \mid \mu, \sigma)$ is the likelihood function

$f(\mu,\sigma)$ is the prior

C is a normalization constant

D is the research data

The likelihood function is

$$L(D \mid \mu,\sigma) = \frac{1}{(2\pi)^{N2} \mid \Sigma \mid^{12}} \exp[-Z^{T} \Sigma^{-1} Z] \tag{3.3.5}$$

where N is the number of data in the implication
and

$$Z_k = \frac{1}{\sigma}\left(Y_{tk} - \mu t_k - \frac{\sigma^2}{2} t_k \right); k = 1,\ldots, N \tag{3.3.6}$$

And

$$\Sigma_{ij} = E[B_{ti} B_{tj}] = \frac{1}{2}(t_i + t_j - \mid t_i - t_j \mid); i.j = 1,\ldots, N \tag{3.3.7}$$

Markov Chain Monte Carlo (MCMC) method used to determine the posterior (3.3.4). Given

$\theta = [\theta_1 \ldots \theta_p]$ is a vector of the parameters $\theta_j; j = 1,\ldots, p$

where

p is the total number of the parameters that were calculated.

The method runs as follow

1) Draw an individual θ_{l+1} from $q(\theta_{l+1} \mid \theta_l; l = 1,\ldots, M)$
2) Accept the individual θ_{l+1} with the probability of $\alpha(\theta_l, \theta_{l+1})$; otherwise set $\theta_{l+1} = \theta_t$

Therefore, the probability of acceptance is

$$\alpha(\theta_l, \theta_{l+1}) = \min\left(\frac{\pi(\theta_l + 1) / q(\theta_{l+1} \mid \theta_l)}{\pi(\theta_l) / q(\theta_l \mid \theta_{l+1})}, l \right) \tag{3.3.8}$$

The independent density, regard a probability density function (PDF), the next individual does not depend on the current one: $q(\theta_{l+1}) = q(\theta_{l+1})$

Refer to the independent density, the probability of acceptance is

1) Draw an individual θ_{l+1} from $q(\theta_{l+1})$

2) Accept the individual θ_{l+1} with the probability of $\alpha(\theta_l,\theta_{l+1})$.

and the probability of acceptance is

$$\alpha(\theta_l,\theta_{l+1}) = \min\left(\frac{\pi(\theta_l+1)/q(\theta_{l+1})}{\pi(\theta_l)/q(\theta_l)}, l\right) \tag{3.3.9}$$

For forecasting the ridership growth rate based on the estimates of model parameters, Monte Carlo simulation is applied to obtain the sample paths.

Recall, one of the SBM properties as follow

$$dB_t = B_{t+dt} - B_t \tag{3.3.10}$$

Then, dB_t is normal distribution with mean zero and variance dt, $dB_t \sim N(0, dt)$. Consequently,

$$B_{t_{k+1}} - B_{t_k} = \sqrt{dt_k} Z_{t_k} \tag{3.3.11}$$

$$Z_{t_k} \sim N(0,1) \tag{3.3.12}$$

Using the Euler Scheme, one has

$$S_{t_{k+1}} = S_{t_k} + S_{t_k}[\mu(t_{k+1} - t_k) + \sigma(B_{k+1} - B_k)]$$

$$k = 0, \ldots, n-1 \tag{3.3.13}$$