Comparison of Empirical Bayes Method and Before-After Study Method

Park Namjune [a], Lee Young Ihn[b]

a,b Graduate School of Environmental studies, Seoul National University, Seoul, 151-742, Korea
a E-mail: namjune1@hanmail.net
b E-mail: yilee@snu.ac.kr

Abstract: Road safety audit is the preventive enhancement strategy for safety. It gets rid of beforehand the potential factor of a traffic accident in the stage of road planning and design and it evaluates the appropriation for road geometric structure or safety facility to prevent traffic accident in the stage of operation after the construction.
Before-After Study method may be overestimated, and considerably dropped about accuracy and explanatory power. So this is a trend that does not use. On the contrary, Empirical Bayes Method has high accuracy and reliability, and has explanatory power. However It have Downside like a narrow analysis range. In our future, Empirical Bayes Method needs more wide variety of analysis. Methodology for him to find is one of the challenges facing us. This paper urged to take advantage of a variety of empirical Bayes methods. In order to support this, paper show a concrete example of utilizing Before-After study Method.

Keywords: Empirical Bayes Method, Before-After Study, SPF, T-Test

1. INTRODUCTION

Road safety audit is the preventive enhancement strategy for safety. : it gets rid of beforehand the potential factor of a traffic accident in the stage of road planning and design and it evaluates the appropriation for road geometric structure or safety facility to prevent traffic accident in the stage of operation after the construction. Since this strategy is introduced to our country in the early 2000s, various projects have been processed and it was legislated recently. And now, the evaluation of past project for its continuation is needed.

The Empirical Bayes Method is use of the evaluation of traffic safety policy, the prediction of traffic accidents and standard technique. Similar road with analysis object road is able to estimate the risk of the analysis object road’s accidents. The Empirical Bayes Method is mainly used in the current domestic, but Before-After Study Method was used in the past. Before-After Study Method cannot overcome the bias by regression toward the mean. It just can get solved by The Empirical Bayes Method. For this reason, Korea more uses Empirical Bayes Method than Before-After Study Method.

The Empirical Bayes Method has the uppermost limit on research methodology. On the contrary, Before-After Study Method had vastly utilized in the past. So Before-After Study Method has an Immense research methodology. In this sense, This paper suggest that The Empirical Bayes Method should be used extensively like Before-After Study Method. So we conducted an analysis using Before-After Study Method, and statistical analysis is one of them.

In this paper, major study object is ‘urge’ about diversify research methodology when we use Empirical Bayes Method. We took the following example in the pursuit of this purpose. The spatial extent of the study is 14 section installed with Skid Proof Pavement in

2009. So accident pre-period is 2007~2008, post-period is 2010~2011. There are temporal extent in this example. Empirical Bayes Method uses reference group, for that reason it needs similar design road’ s data. For example, there are geometric design of roadway, number of vehicle and number of accidents.

2. Literature review

Kim-Tae-Young did comparative analysis using Empirical Bayes Method. This paper is used Safety Effectiveness Factor. This paper found accident reduction effect by traffic signal violation enforcement service.

Park-Min-Ho studied the impact of traffic enforcement on reducing traffic accidents by Empirical Bayes Method. This paper found traffic enforcement is an effective policy.

Mun-Seung-Ra did effectiveness assessment using Empirical Bayes Method. Accident effectiveness is analyzed using Safety Effectiveness Factor in this paper.

Persaud et al. did effectiveness assessment when switch from intersection to revolving intersection. In consequence accidents declined from 60% to 40%.

3. Effectiveness assessment Methodology

3.1. Empirical Bayes Method

In 1997, Hauer found Empirical Bayes Method. This Methodology compare the number of realistic accidents number and reference gorup’s accidents number. Reference group is similar in realistic design and realistic number of vehicle. The Empirical Bayes Method eliminates RTM(regression to the mean), so it can find accurate analysis result. For this reason most of researcher use this method.

The Empirical Bayes Method has two types of model. There are called gamma distribution and poisson distribution. Gamma distribution use prior distribution, but poisson distribution use likelihood. When estimating Poisson parameter(λ), we can find uncertainty. In this model, the number of traffic accidents observation is y.

$$f(\lambda \mid y) = \frac{f(y \mid \lambda) \bullet f(\lambda)}{f(y)} \tag{1}$$

The negative binomial distribution observed y and the likelihood of the joint distribution of the prior distribution of the Poisson parameters from the number of accidents can be derived.

$$f(y_i \mid x_i) = \frac{\Gamma(y_i + \alpha)}{\Gamma(y_i + 1)\Gamma(\alpha)} v_i^{\alpha} (1 - v_i)^{y_i} \tag{2}$$

where

y_i : observed traffic accident,

v_i : exponential average function

$\Gamma(*)$: gamma function.

Use the appropriate model of them, and then calculates Safety Performance Function and Over-dispersion Parameter. We needs Over-dispersion Parameter when difference

predicted number of accident between realistic number of accident.

$$\varphi = \Phi_i \times SPF_i^r \tag{3}$$

where

Φ : over-dispersion parameter,
SPF : safety performance function.

$$w = \frac{1}{(1+E(k))/\varphi}$$
$$\pi = E(k \mid K) = \omega E(k) + (1-\omega)k \tag{4}$$

where

$E(k)$: predicted number of accident,
k : realistic number of accident,
$E(k)$: safety performance function, SPF,
ω : weighted value.

3.2. Before-After study method

Shen, Gan explained Before-After Study Method could classified into the four ways. There are four way of effective analysis.

1. Simple comparison
2. one to one matching with yoked comparison
3. before-and-after study with comparision on group
4. C.G. Method

Simple comparison is easy to calculate and understand, but it has regression to the bias. one to one matching with yoked comparison can’ t use zero point. (number of accident: zero) This method has regression to the bias, too. On the contrary, The Empirical Bayes Method eliminates RTM(regression to the mean), so it can find accurate analysis result. For this reason most of researcher use this method. However it has disadvantage of requiring a lot of data sample.

$$CRF = \frac{(N_b - N_a)}{N_b} = 1 - \frac{N_a}{N_b} \tag{4}$$

where

N_b : number of accident (before),
N_a : number of accident (after).

4. Result

4.1. Effective assessment of Empirical Bayes Method

When we use the Empirical Bayes Method, we have to follow the prearranged procedure. EB Method's procedure summary is shown <Table 2>.

<Table 2> Empirical Bayes Method's procedure

Procedure #1. SPF model construction using Limdep
Procedure #2. Calculate E(k): E(k) used for calculating "safety performance function(SPF)".
Procedure #3. Calculate weighted value(w) using $E(k)$,
Procedure #4. Calculate EB value, E(k/K)
Procedure #5. Calculate accident rate of change(En) using E(k/K), number of accident(L)

First we have to find safety performance function by using LIMDEP. LIMDEP takes the form of an econometrics studio. Analysis of a data set is done interactively in a set of windows. Program control may be from a 'script' or in an unstructured session of instructions and manipulations. The program is designed to allow easy setup of data for estimation, specification of different forms of the models, experimentation with different specifications, hypothesis testing, analysis of data and model results and construction of special procedures and estimators.

So we can estimate value of $E(k)$, so we can find weighted value(w). And then we can get odds ratio(θ). So we can calculate rate of accident reduction. At last using calculate rate of accident reduction, compare and analysis with existing data.

$$\pi = E(k \mid K) = \omega E(k) + (1-\omega)k \tag{5}$$

where

$E(k)$: predicted number of accident,
k : realistic number of accident,
ω : weighted value

<Table 3> and <Table 4> mean road's information. (Observed road data: preexistence data, Reference group data: processed data)

<Table 3> Observed road

	road form	construction date	traffic volume	number of accident	
			before	before	after
Observed Road #1	single	2009	12,023	2	0
Observed Road #2	single	2009	28,544	4	1
Observed Road #3	single	2009	29,856	5	3
Observed Road #4	single	2009	5,952	2	0
Observed Road #5	inter	2009	44,927	6	1
Observed Road #6	inter	2009	41,183	6	5
Observed Road #7	inter	2009	23,718	3	2
Observed Road #8	inter	2009	39,840	5	1

Observed Road #9	inter	2009	4,736	2	1
Observed Road #10	inter	2009	41,183	12	8
Observed Road #11	inter	2009	15,288	4	3
Observed Road #12	inter	2009	22,480	4	0
Observed Road #13	inter	2009	62,401	18	16
Observed Road #14	inter	2009	43,031	17	13

number of accident(before) = 2007'~2008'

number of accident(after) = 2010'~2011'

<Table 4> Reference group

	road form	traffic volume	number of accident	accident rate (%)
Reference group #1	single	12,969	3	0.634
Reference group #2	single	28,544	5	0.48
Reference group #3	single	29,078	5	0.471
Reference group #4	single	4,608	2	1.189
Reference group #5	inter	50,544	7	0.379
Reference group #6	inter	47,478	7	0.404
Reference group #7	inter	24,701	3	0.333
Reference group #8	inter	41,919	5	0.327
Reference group #9	inter	5,312	2	1.032
Reference group #10	inter	37,253	12	0.883
Reference group #11	inter	14,539	4	0.754
Reference group #12	inter	21,024	4	0.521
Reference group #13	inter	59,966	16	0.731
Reference group #14	inter	52,540	19	0.991

To model building, we need to estimate the safety performance function by using Limdep. For this, several of the independent variables have to be considered.

<Table 5> Regression model

variables	regression model type	standard error	coefficient significance probability	significance probability
traffic volume	Poisson Regression	0.663	0.374	0.011*
	Negative Binomial Regression	0.289	0.042*	0.000**

road type	Poisson Regression	0.238	0.531	0.181
	Negative Binomial Regression	0.476	0.921	0.370
traffic volume, road type	Poisson Regression	0.554	0.579	0.033*
	Negative Binomial Regression	0.385	0.700	0.120

(p<0.01:**, p<0.05:*)

Traffic is independent variables used to predict the traffic accidents. In this table, we can find significance probability is exist when variable is traffic volume. Poisson regression have 0.011* compared with negative binomial regression has 0.000**

<Table 6> Estimate coefficient

		Poisson Regression	Negative Binomial Regression
coefficient	Constant	0.589	0.586
	traffic volume	0.364D-04	0.365D-04
Standard Error	Constant	0.663	0.289
	traffic volume	0.142D-04	0.656D-05
b/Standard Error	Constant	0.889	2.030
	traffic volume	2.554	5.565
significance probable	Constant	0.374	0.042*
	traffic volume	0.011*	0.000**
Mean X	-	30748.214	

In this study, we considered only three variables. (traffic volume, road type, traffic)
SPF equation is derived as follows.

$$Y = 0.0000656X - 0.586 \quad (5)$$

Progress by taking advantage of this expression, calculated according to the order in Table 2 to obtain the expected number of accidents and dispersion coefficients and weights, and target road and directly comparable. Improve it through to the final are also available. Table 7 below appear two sections with no improvement and worsening, except for one section in 11 sections was reduced from an average of 30% of the incidence of traffic accidents. It was noticeable improvement in traffic accidents slip-resistant packaging.

<Table 7> Expected number of accident and Change rate(%)

	SPF	overdispersion parameter(¢)	weighted value(w)	realistic number of accident	expected number of accident	change rate(%)
1	0.2	0.036	0.030	2	2	0.0
2	1.3	0.233	0.101	4	3	-25.0
3	1.3	0.233	0.101	5	3	-40.0
4	-0.3	-0.054	-0.077	2	2	0.0
5	2.7	0.483	0.131	6	4	-33.3
6	2.5	0.448	0.128	6	4	-33.3
7	1	0.179	0.090	3	2	-33.3
8	2.2	0.394	0.123	5	3	-40.0
9	-0.2	-0.036	-0.045	2	3	50.0
10	1.9	0.340	0.117	12	9	-25.0
11	0.4	0.072	0.051	4	3	-25.0
12	0.8	0.143	0.080	4	2	-50.0
13	3.3	0.591	0.137	18	14	-22.2
14	2.9	0.519	0.133	17	13	-23.5

4.2. Effective assessment(Before-After study Method)

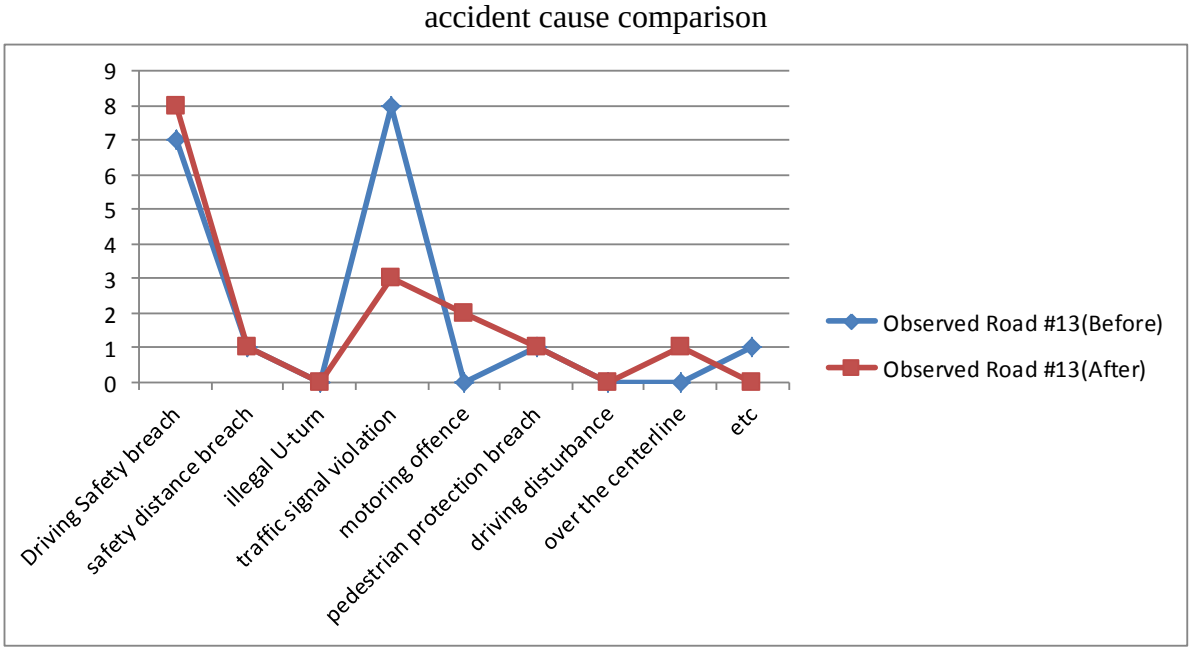
Certain conditions, to changes in the value of the before and after to compare statistical methods are used to analyze the effect of changing conditions. Typically, It is used in the statistical material is refined through controlled experiments. It is hard that traffic data collects due to wide range of variable conditions.

<Table 10> Observed road's data #3

	intersection	cause of accident								
		Driving Safety breach	safety distance breach	illegal U-turn	traffic signal violation	motoring offence	pedestrian protection breach	driving disturbance	over the centerline	etc
B	Observed Road #1	1	0	0	0	0	0	0	1	0
E	Observed Road #2	2	0	0	0	0	0	0	2	0
F	Observed Road #3	5	0	0	0	0	0	0	0	0

O R E	Observed Road #4	2	0	0	0	0	0	0	0	0
	Observed Road #5	5	1	0	0	0	0	0	0	0
	Observed Road #6	5	1	0	0	0	0	0	0	0
	Observed Road #7	3	0	0	0	0	0	0	0	0
	Observed Road #8	2	0	0	3	0	0	0	0	0
	Observed Road #9	2	0	0	0	0	0	0	0	0
	Observed Road #10	6	0	0	4	0	0	0	2	0
	Observed Road #11	2	0	0	1	1	0	0	0	0
	Observed Road #12	2	0	0	1	1	0	0	0	0
	Observed Road #13	7	1	0	8	0	1	0	0	1
	Observed Road #14	7	5	0	5	0	0	0	1	0
A F T E R	Observed Road #1	0	0	0	0	0	0	0	0	0
	Observed Road #2	0	0	0	0	0	0	0	1	0
	Observed Road #3	2	1	0	0	0	0	0	0	0
	Observed Road #4	0	0	0	0	0	0	0	0	0
	Observed Road #5	0	0	0	0	0	0	0	0	1
	Observed Road #6	4	0	1	0	0	0	0	0	0
	Observed Road #7	3	0	0	0	0	0	0	0	0
	Observed Road #8	1	0	0	0	0	0	0	0	0
	Observed Road #9	1	0	0	0	0	0	0	0	0
	Observed Road #10	3	0	0	2	2	0	1	0	0
	Observed Road #11	2	0	0	1	0	0	0	0	0
	Observed Road #12	0	0	0	0	0	0	0	0	0
	Observed Road #13	8	1	0	3	2	1	0	1	0
	Observed Road #14	8	3	0	2	0	0	0	0	0

4.3. Analysis result



There are comparing accident cause(before) with after's. We can get a direct comparison of numbers, as well as a comparison with the average value. Thus, comparative analysis in the pre-and post-status alone simply can get a lot of information.
The same sample t-test was conducted to verify the various differences.

<Table 12> T-test (sample #1)

		mean	standard deviation	t	p
snow	before	.36	.497	2.687	.019*
	after	.00	.000		
rain	before	.79	1.122	1.000	.336
	after	.57	.938		

(p<0.01:**, p<0.05:*)

<Table 13> T-test (sample #2)

		mean	standard deviation	t	p
rear-end (stop)	before	.57	.852	1.749	.104
	after	.29	.469		
rear-end (driving)	before	.79	1.051	2.857	.013*
	after	.14	.363		

(p<0.01:**, p<0.05:*)

<Table 14> T-test (sample #3)

		mean	standard deviation	t	p
death rate	before	.106	.182	2.017	.049*
	after	.01	.021		

<Table 12> compare the difference about the number of accidents in the rain and snow and their mean. we can check skid proof pavement makes difference in snow. Table 13 shows comparison rear-end(stop) between rear-end(driving). We can get significance probability when rear-end(driving)(p<.05). So we can insist proof pavement makes difference in rear-end(driving). Difference according to the average of the various pre-and post-analysis can be used to verify. These are, respectively, showing a significant difference p <.05 by becoming that can be described statistically.

<Table 15> Improvement(%)

	Empirical Bayes Method			Before-After Study		
	before	after	Improvement (%)	before	after	Improvement (%)
Reference group #1	2	2	0	2	0	-100
Reference group #2	4	3	-25	4	1	-75
Reference group #3	5	3	-40	5	3	-40
Reference group #4	2	2	0	2	0	-100
Reference group #5	6	4	-33.3	6	1	-83.3

Reference group #6	6	4	-33.3	6	5	-16.7
Reference group #7	3	2	-33.3	3	2	-33.3
Reference group #8	5	3	-40	5	1	-80
Reference group #9	2	3	50	2	1	-50
Reference group #10	12	9	-25	12	8	-33.3
Reference group #11	4	3	-25	4	3	-25
Reference group #12	4	2	-50	4	0	-100
Reference group #13	18	14	-22.2	18	16	-11.1
Reference group #14	17	13	-23.5	17	13	-23.5

Reference group = The end result's object

Look at the results of the above analysis, results differ depending on the analytical method. The Empirical Bayes Method eliminates RTM(regression to the mean), so it can find accurate analysis result. For this reason most of researcher use this method.

<Table 16> Result comparison

	before	after	Improvement(%)
Empirical Bayes Method	2	2	0
Before-After Study	2	0	-100

Look at Table 16, the frequency of occurrence of an accident the past 2 actually slip-resistant packaging after 0 review accident quite a traffic accident. Nevertheless, It can not say it was reduced. Because in Empirical Bayes Method before and after that if the same number of accidents. It means non-human variable geometric factor due to the actual number of accidents had reportedly reduced the number of accidents.

5. Conclusion

This paper analyzed a total of 14 sections conducted skid proof pavement in 2009 for compare Empirical Bayes Method between Before-After Study. Before-After study analysis results showed that skid proof pavement to reduce traffic accidents. By contrast, the empirical Bayes method appeared to affect only section 11 to reduce traffic accidents. The Empirical Bayes Method eliminates RTM(regression to the mean), so it can find accurate analysis result. For this reason most of researcher use this method. Before-After Study method may be overestimated, and considerably dropped about accuracy and explanatory power. So this is a trend that does not use. On the contrary, Empirical Bayes Method has high accuracy and reliability, and has explanatory power. However It have Downside like a narrow analysis range. In our future, Empirical Bayes Method needs more wide variety of analysis. Methodology for him to find is one of the challenges facing us. This paper urged to take

advantage of a variety of empirical Bayes methods. In order to support this, paper show a concrete example of utilizing Before-After study Method.

Reference

Harwood, D., et al., "Safety Effectiveness of Intersection Left- and Right-Turn Lanes", Transportation Research Record 1840, pp.131-139, 2003

Hauer, E., "Observational Before-After Studies in Road Safety", Pergamon. Oxford, UK, 1997

Persaud, B., et al., "Observational Before -After Study of the Safety Effect of U.S. Roundabout Conversions Using the Empirical Bayes Method", Transportation Research Record 1751, pp.1-8, 2001

Persaud, B., et al., "Safety Evaluation of Permanent Raised Snow-Plowable Pavement Markers", Transportation Research Record 1897, pp.148-155, 2004