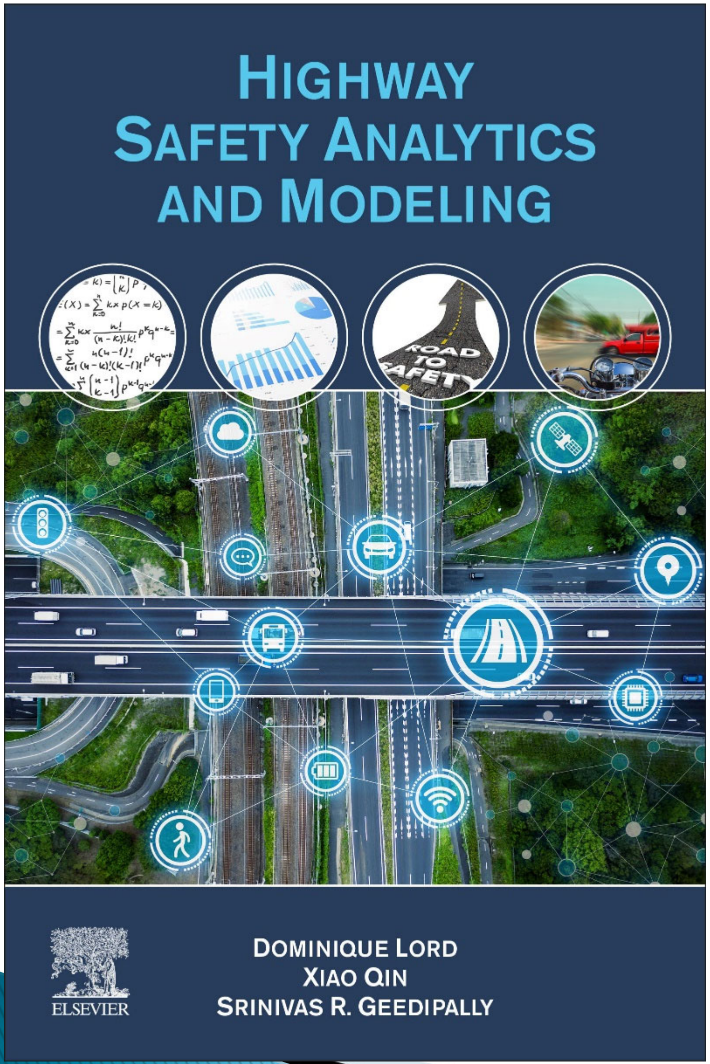
Before-After Studies

Part 1

December 22, 2021

Instructor: Dominique Lord
Texas A&M University
d-lord@tamu.edu

Textbook



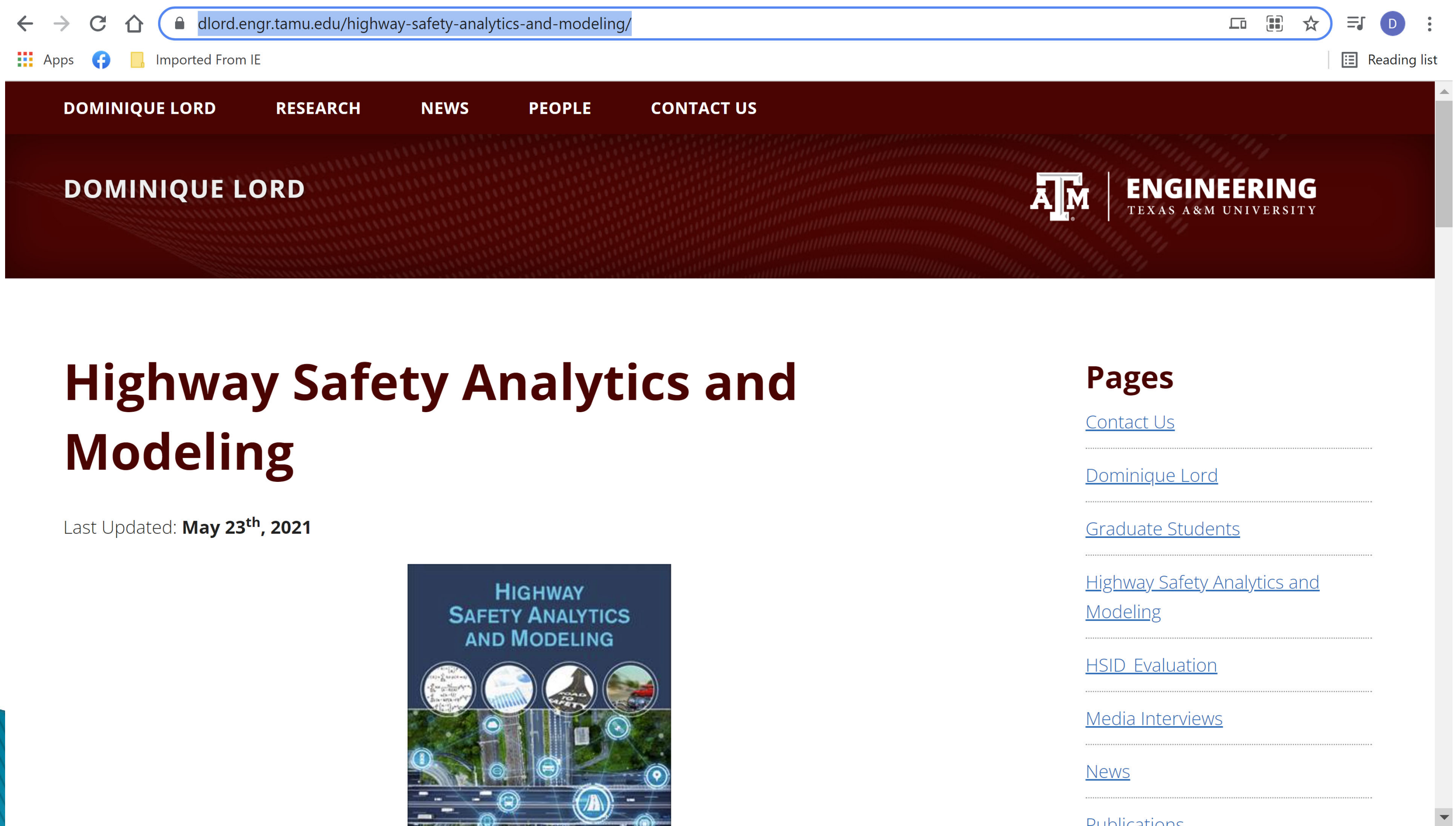
The material presented in this series of lectures are taken from this textbook and other sources based on lectures given by the authors.

The textbook is available on Amazon and the Elsevier website below among other places.

https://www.elsevier.com/books/highway-safety-analytics-and-modeling/lord/978-0-12-816818-9

Textbook

Datasets for examples and updates/corrections can be find in the following link:
https://dlord.engr.tamu.edu/highway-safety-analytics-and-modeling/



BEFORE-AFTER STUDIES

- Why before-after studies?
  - As opposed to cross-sectional studies, the before-after study has lower within-subject variability (i.e., the variation associated with multiple measurements observed over time for one subject)
  - Hence, the analyst has a better control about the effects of an intervention on safety and crash risk.
- Types of studies
  - Naïve or Simple before-after studies
  - Before-after studies with control group
  - Empirical Bayes approach (control group)
  - Full Bayes
- Important Issues to account for
  - Regression-to-the-mean
  - Selection bias

Regression-to-the-mean

The regression-to-the-mean phenomenon is commonly associated to random events and consists of the general tendency of extreme values to regress to median values.

Crash data have been shown to exhibit this characteristic.

Consequently, the RTM can have a significant effect on the evaluation of treatments, and this occurs exclusively for before-after studies.

When the crash frequency is abnormally high during in a given period, it tends to decrease during the subsequent period and draws closer to the site's long-term average (or vice-versa).

Regression-to-the-mean

Table 11.2. Accident count at 1142 intersections - 1974/1975.

1 Number of Intersections n(K)	2 Number of Accidents per intersection in 1974 K	3 Average Number of Accidents per intersection in 1975 avg(K)
553	0	0.54
296	1	0.97
144	2	1.53
65	3	1.97
31	4	2.10
21	5	3.24
9	6	5.67
13	7	4.69
5	8	3.80
2	9*	6.50

* In addition, 2 intersections had 13 accidents, one had 16.

Regression-to-the-mean

Table 11.3. Accident counts at 1072 intersections with up to 9 accidents in 1974-77.

1 Number of Intersections n(K)	2 Accidents per Intersection in 1974-1976 K	3 Average Accidents per Intersection in 1977 avg(K)
256	0	0.25
218	1	0.55
173	2	0.70
121	3	1.04
97	4	1.08
70	5	1.33
54	6	1.56
32	7	2.25
29	8	1.62
22	9	2.50

Regression-to-the-mean

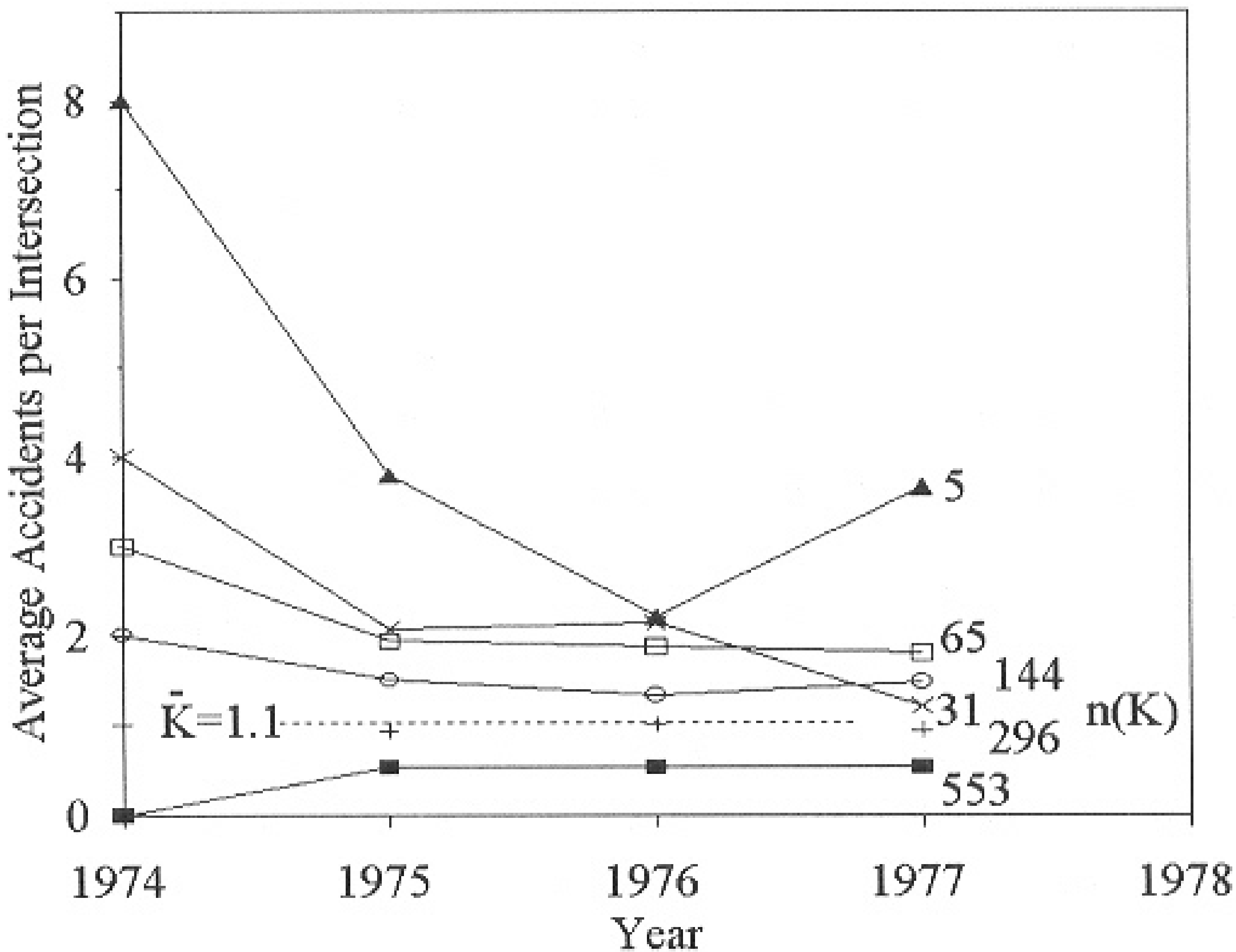


Figure 11.1. How accident counts regress to the mean.

Regression-to-the-mean

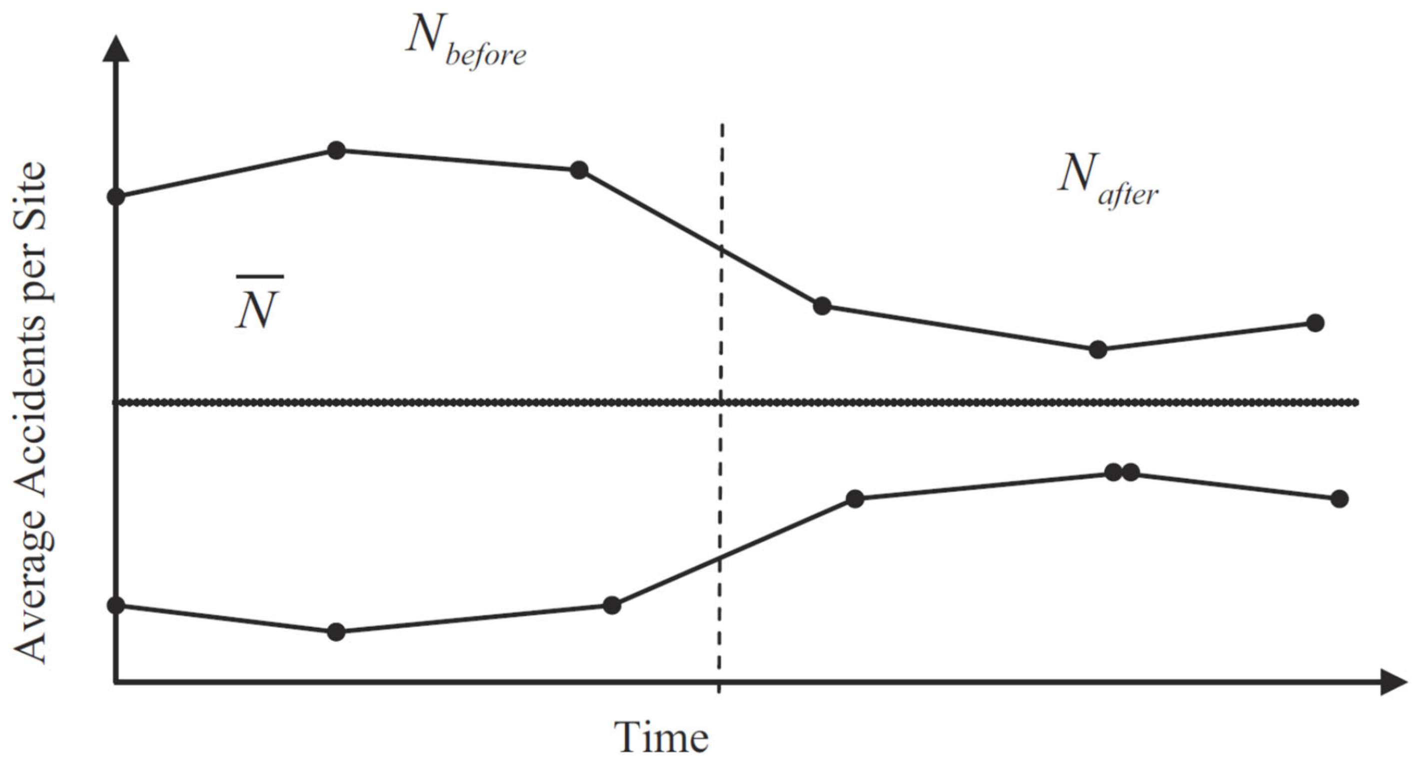


FIGURE 7.1 Representation of the regression-to-the-mean (Lord and Kuo, 2012).

Regression-to-the-mean

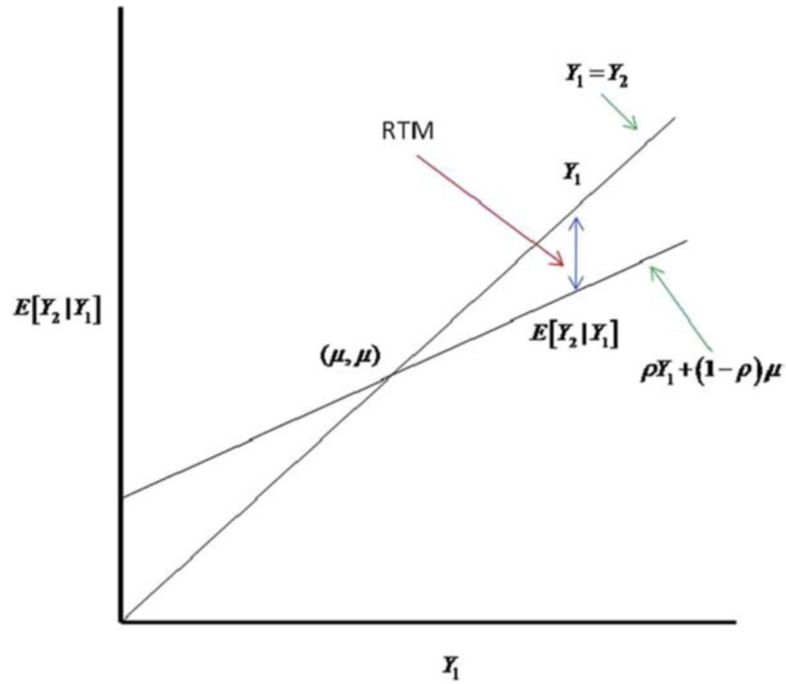


FIGURE 7.2 Relationship between $E[Y_2|Y_1]$ and Y_1 (Lord and Kuo, 2012).

$$E[Y_2|Y_1] = \rho Y_1 + (1-\rho)\mu$$

When the correlation coefficient is equal to 1, no RTM exist as $E[Y_2|Y_1]=Y_1$. On the other hand, when the correlation coefficient is not equal to 1, RTM is observed in the data. Smaller values of r are associated with larger RTM effects because $E[Y_2|Y_1]$ is closer to μ and farther away from Y_1.

Site Selection Bias

Site selection bias refers to sites that are solely selected based on high crash count experience.

This bias obviously affects the outcome of before-after studies. Ideally, the evaluation of different alternatives should be performed using a randomized trial (a mix of sites with different long-term averages).

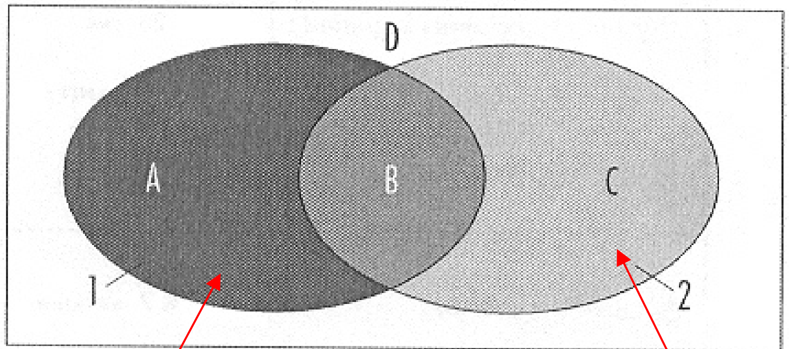
Unfortunately, sites used for evaluating different alternatives are often selected based on the crash counts.

For instance, Warrant 7 from the MUTCD indicates that traffic signal control should be contemplated on a site where more than 5 crashes occurred in a 12-month period.

This Warrant is applicable only if this control can be used to reduce the number of collisions at that given site. (Discussed further later).

Site Selection Bias

Figure 5-A5 Selection bias



A: unsafe sites undetected
B: unsafe sites detected as being hazardous
C: normal sites detected as being hazardous
D: normal sites undetected

High Long-term average

High crash count

Site Selection Bias

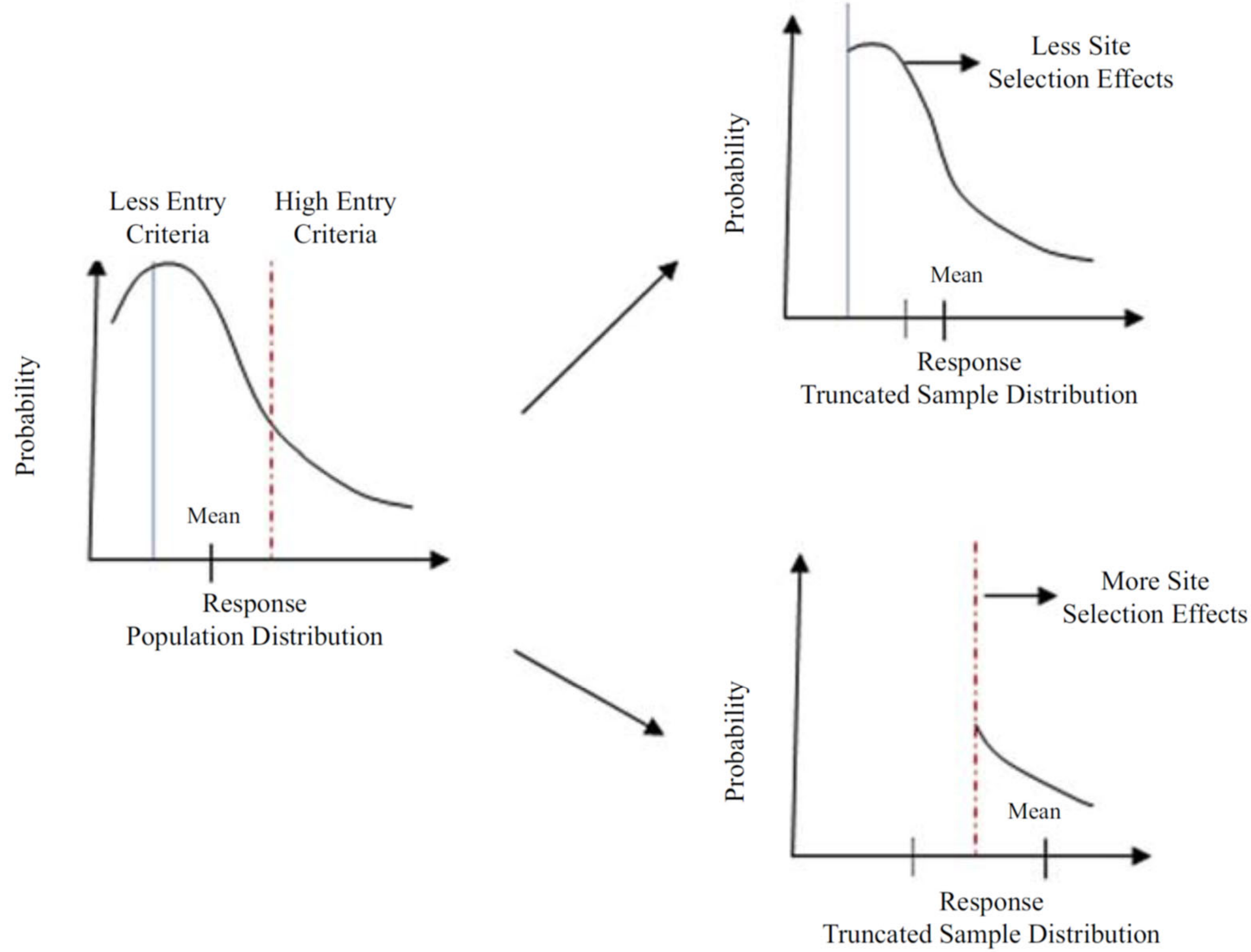


FIGURE 7.3 The population distribution for complete and truncated samples (Lord and Kuo, 2012).

Before-After Studies

There are many variants of Before-After studies. They can be regrouped under two tasks:

1. Predict what would have been the safety of an entity in the "after" period, had the treatment not been applied, and
2. Estimate what the safety of the treated entity in the after period was.

An entity is a general term used to designate a road section, intersection, ramp, driver, etc.

The analysis can be divided into four basic steps.

Before-After Studies

First, we need to define the notation that will be used for performing the two tasks at hand.

Let:

π be the expected number of target crashes of a specific entity in an after period would have been had it not been treated; π is what must be predicted.

λ be the expected number of target crashes of a specific entity in an after period; λ is what must be estimated.

Technically, we should be using ^ on top of the term since those are estimates. However, for simplicity reason, we will not use it.

Before-After Studies

The effect of a treatment is judge by comparing π and λ . The two comparisons we are usually interested are the following:

$\delta = \pi - \lambda$ the reduction in the after period of the expected number of target crashes (by kind and severity).

$\theta = \lambda / \pi$ the ratio of what was the treatment to what it would have been without the treatment; this is defined as the index of effectiveness.

Before-After Studies

Example: suppose that a treatment has been implemented in 1992. Now, suppose that if the treatment would not have been implemented, one would have expected 360.6 crashes (or $\pi = 360.6$) in 1995 and 1996. We know that the estimated number of crashes that occurred in 1995 and 1996 was 295 (or $\lambda = 295$). Estimate the change in safety:

Before-After Studies

Example: suppose that a treatment has been implemented in 1992. Now, suppose that if the treatment would not have been implemented, one would have expected 360.6 crashes (or $\pi = 360.6$) in 1995 and 1996. We know that the estimated number of crashes that occurred in 1995 and 1996 was 295 (or $\lambda = 295$). Estimate the change in safety:

$$\delta = \pi - \lambda$$

$$\delta = 360.6 - 295.0$$

$$\delta = 65.6$$

A reduction of 65.6 in target crashes

$$\theta = \lambda / \pi$$

$$\theta = 295.0 / 360.6$$

$$\theta = 0.82$$

A 18% reduction in target crashes

Before-After Studies

Using the assumption that the crash counts are Poisson distributed, the variance of π and λ are given as follows:

$$Var\{\lambda\} = \lambda$$

Usually, it is assumed that observed crashes are Poisson distributed for any given site.

$$Var\{\pi\} =$$

Will depend on the method used for predicting the value (e.g., Poisson-gamma).

Before-After Studies

The estimation of the safety of a treatment is done through a 4-step process. This step is done for each entity.

STEP 1: Estimate λ and predict π . There are many ways to estimate or predict these values. Some will be shown in this lecture (and textbook).

STEP 2: Estimate $Var(\lambda)$ and $Var(\pi)$. These estimates depend on the methods chosen. Often, λ is assumed to be Poisson distributed, thus $Var(\lambda) = \lambda$.

If a statistical model is used:

$$Var(\pi) = \frac{\pi^2}{\phi}$$

Same as μ for Poisson or Poisson-gamma model

Before-After Studies

The estimation of the safety of a treatment is done through a 4-step process.

STEP 3: Estimate δ and θ using λ and π from STEP1 and $Var(\pi)$ from STEP 2.

$$\delta = \pi - \lambda$$

$$\theta = \frac{\lambda}{\pi\left[1 + Var\{\pi\} / \pi^2\right]}$$

Before-After Studies

The estimation of the safety of a treatment is done through a 4-step process.

STEP 3: Estimate δ and θ using λ and π from STEP1 and $Var(\pi)$ from STEP 2.

$$\delta = \pi - \lambda$$

$$\theta = \frac{\lambda}{\pi\left[1 + Var\{\pi\} / \pi^2\right]}$$

Correction factor when fewer than 500 observations are used.

Before-After Studies

The estimation of the safety of a treatment is done through a 4-step process.

STEP 4: Estimate $Var\{\delta\}$ and $Var\{\theta\}$.

$$Var\{\delta\} = Var\{\pi\} + Var\{\lambda\}$$

$$Var\{\theta\} \cong \frac{\theta^2 \left[\left(Var\{\lambda\} / \lambda^2 \right) + \left(Var\{\pi\} / \pi^2 \right) \right]}{\left[1 + Var\{\pi\} / \pi^2 \right]^2}$$

Before-After Studies

The safety estimation of a treatment is done through a 4-step process.

When you have more than one site:

$$\lambda = \sum \lambda_i$$

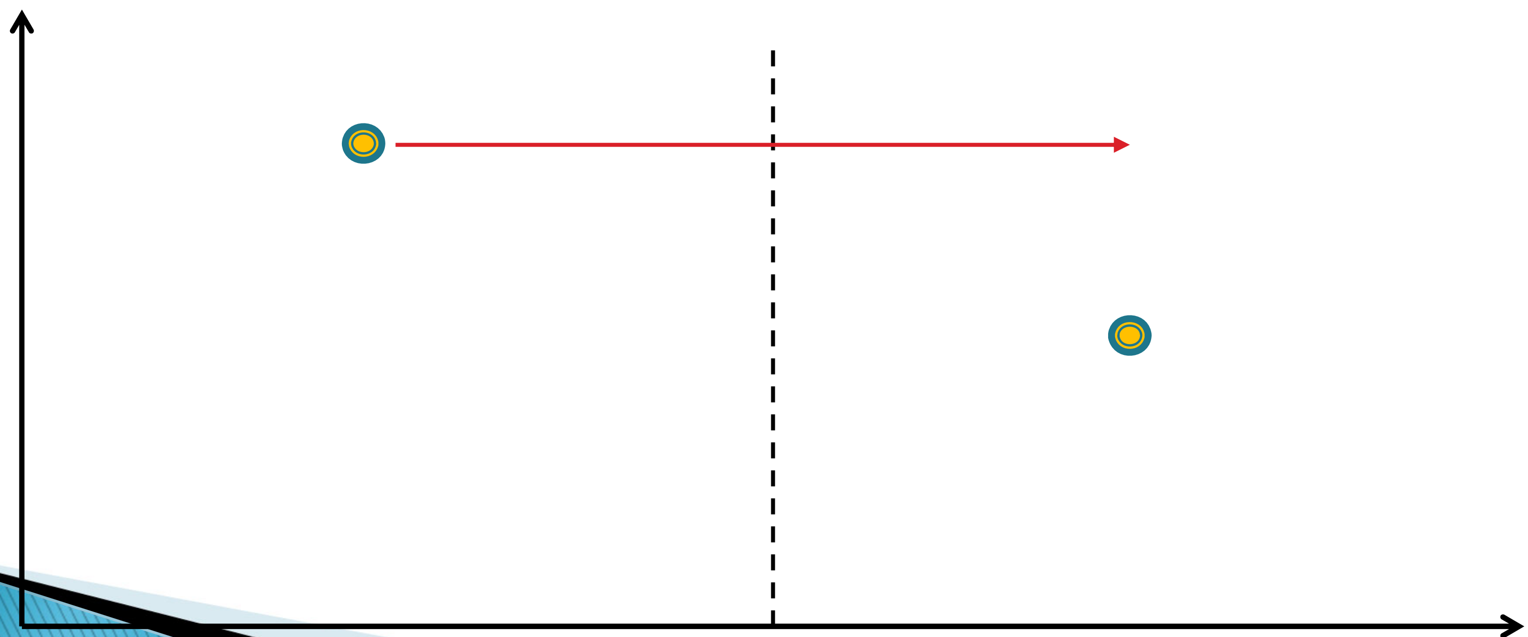
$$\pi = \sum \pi_i$$

$$Var\{\lambda\} = \sum Var\{\lambda_i\}$$

$$Var\{\pi\} = \sum Var\{\pi_i\}$$

Naïve/Simple Before-After Studies

In its simplest form, an observational before-after study consists of comparing the counts occurring in the before period to its count in the after period. The term naïve stands for the fact that counts in the before period are used as predictor of the expected crashes occurring in the after period.



Time

Naïve Before-After Studies

Limitations:

- Traffic, weather, road user behavior, vehicle fleet changes over time
- Besides the treatment of interest, various other treatments or programs may be implemented at the same time
- PDO counts may change over time
- The probability of crashes being reported may be changing over time
- Entities selected for treatment may be selected because of unusual crash experience (selection bias)
- Does not account for RTM

Naïve Before-After Studies

Disclaimer 1:

"The noted change in safety reflects not only the effect of …(name of treatment)… but the effect of factors such as traffic, weather, vehicle fleet, driver behavior, cost of vehicle repairs, inclination to report crashes and so on. It is not known what part of the changes can be attributed to …(name of treatment)… and part due to other influences"

Disclaimer 2:

"The noted change in safety may be in part due to the spontaneous regression-to-the-mean and not due to…(name of treatment)."

Naïve Before-After Studies

Let λ(1), λ(2), … λ (n) represent the crash counts occurring on site j for the before period.

Let κ(1), κ(2), … κ (n) represent the crash counts occurring on site j for the after period.

Let the "ratio of duration" to be:

$$r_{dj} = \frac{\text{Duration of after period for entity j}}{\text{Duration of before period for entity j}}$$

Naïve Before-After Studies

STEP 1 & STEP 2

Estimates of Coefficients

$$\lambda = \sum \lambda_j$$

$$\pi = \sum r_{dj} \kappa_j$$

$$\pi = r_d \sum \kappa_j$$

Estimates of Variances

$$Var\{\lambda\} = \sum \lambda_j$$

$$Var\{\pi\} = \sum r_{dj}^2 \kappa_j$$

$$Var\{\pi\} = r_d^2 \sum \kappa_j$$

If r_d is the same

Naïve Before-After Studies

STEP 3 & STEP 4

$$Var\{\delta\} = Var\{\pi\} + Var\{\lambda\}$$

$$\theta = \frac{\lambda}{\pi\left[1 + Var\{\pi\} / \pi^2\right]}$$

$$Var\{\theta\} \cong \frac{\theta^2\left[\left(Var\{\lambda\}/\lambda^2\right) + \left(Var\{\pi\}/\pi^2\right)\right]}{\left[1 + Var\{\pi\}/\pi^2\right]^2}$$

Naïve Before-After Studies

Example R.I.D.E. program*

One year before the program was implemented, there were 173 alcohol-related crashes that occurred in one of the five police districts.

In the year after its implementation, 144 alcohol-related collisions occurred in the same district.

Estimate the change in safety.

*Reduce Impaired Driving Everywhere (Ontario)

Naïve Before-After Studies

Example R.I.D.E. program

One year before the program was implemented, there were 173 alcohol-related crashes that occurred in one of the five police districts.

In the year after its implementation, 144 alcohol-related collisions occurred in the same district.

Estimate the change in safety.

$$\lambda = 144 \qquad Var\{\lambda\} = 144$$

$$\pi = 173 \qquad Var\{\pi\} = 173$$

$$\delta = 173 - 144 = 29 \qquad Var\{\delta\} = 173 + 144 = 317$$

$$\theta = \frac{144/173}{\left(1+\frac{173}{173^2}\right)} = 0.83 \qquad Var\{\theta\} = 0.83^2 \times 0.0126 = 0.0087$$

Naïve Before-After Studies

Exercise 7.1

Using the Before−After Dataset, conduct a before−after study using the naïve method. Assume a treatment was installed at the beginning of Year 6 (i.e., after period is 3 years). The data were collected over 8 years for 15 sites.

First, calculate the ratio r_d

$$r_d = 3/5 = 0.6$$

Estimate π, λ, $Var(\pi)$, $Var(\lambda)$

Site ID	r_d	π	λ	$Var(\pi)$	$Var(\lambda)$
1	0.6	7.2	5	4.3	5
2	0.6	9	9	5.4	9
3	0.6	9.6	5	5.8	5
4	0.6	9.6	5	5.8	5
5	0.6	15.6	9	9.4	9
6	0.6	8.4	5	5.0	5
7	0.6	15	12	9.0	12
8	0.6	11.4	9	6.8	9
9	0.6	11.4	16	6.8	16
10	0.6	10.8	14	6.5	14
11	0.6	17.4	8	10.4	8
12	0.6	15.6	12	9.4	12
13	0.6	3.6	11	2.2	11
14	0.6	8.4	8	5.0	8
15	0.6	18.6	12	11.2	12
Sum		171.6	140	103.0	140

Naïve Before-After Studies

Calculate δ

$$\delta = \pi - \lambda$$
$$\delta = 171.6 - 140$$
$$\delta = 31.6$$

There is a reduction of 31.6 in the expected number of crashes.
Calculate θ (adjust for a sample size below 500)

Naïve Before-After Studies

Exercise 7.1 (*cont'd*)

$$\theta = \frac{\lambda}{\pi[1 + Var(\pi)/\pi^2]}$$

$$\theta = \frac{140}{171.6[1 + 103.0/171.6^2]}$$

$$\theta = 0.82$$

There is a reduction of 18% in the expected number of crashes.
Calculate $Var(\delta)$ and $SD(\delta)$

$$Var(\delta) = Var(\pi) + Var(\lambda)$$
$$Var(\delta) = 103.0 + 140$$
$$Var(\delta) = 243$$

$$SD(\delta) = \sqrt{\delta}$$
$$SD(\delta) = \sqrt{243.0}$$
$$SD(\delta) = 15.6$$

Naïve Before-After Studies

The reduction is $31.6 \pm 1.96 \times 15.6$, which is statistically significant at the 5% level.

Calculate $Var(\theta)$ and $SD(\theta)$

$$Var(\theta) \cong \frac{\theta^2\left[\left(Var(\lambda)/\lambda^2\right)+\left(Var(\pi)/\pi^2\right)\right]}{\left[1+Var(\pi)/\pi^2\right]^2}$$

$$Var(\theta) \cong \frac{0.82^2\left[\left(140/140^2\right)+\left(103.0/171.6^2\right)\right]}{\left[1+103.0/171.6^2\right]^2}$$

$$Var(\theta) \cong 0.007$$

$$SD(\theta) = \sqrt{0.007}$$

$$SD(\theta) = 0.084$$

The reduction is $0.82 \pm 1.96 \times 0.084$, which is statistically significant at the 5% level.

Naïve Before-After Studies

Exercise 7.2

Redo Exercise 7.1, but only include sites that experienced three or more crashes in the before period. Use the last 3 years before and the last 3 years after (i.e., $r_d = 1$).

First, estimate π, λ, $Var(\pi)$, $Var(\lambda)$

Site ID	r_d	π	λ	$Var(\pi)$	$Var(\lambda)$
1					
2	1	10.0	9	10.0	9
3	1	11.0	5	11.0	5
4					
5					
6					
7	1	13.0	12	13.0	12
8					
9	1	11.0	16	11.0	16
10					
11	1	20.0	8	20.0	8
12	1	20.0	12	20.0	12
13					
14					
15	1	22.0	12	22.0	12
Sum		107.0	74	107	74

Naïve Before-After Studies

Calculate δ

$$\delta = 33.0$$

There is a reduction of 33.0 in the expected number of crashes.
Calculate θ (adjust for a sample size below 500)

$$\theta = 0.69$$

There is a reduction of 31% in the expected number of crashes, which is much larger than the reduction observed in Exercise 7.1. This shows the effects of the site selection bias.

Calculate $Var(\delta)$ and $SD(\delta)$

$$Var(\delta) = 181.0$$
$$SD(\delta) = 13.5$$

The reduction is $33.0 \pm 1.96 \times 13.5$, which is statistically significant at the 5% level.

Naïve Before-After Studies

Calculate $Var(\theta)$ and $SD(\theta)$

$$Var(\theta) \cong 0.011$$
$$SD(\theta) = 0.103$$

The reduction is $0.69 \pm 1.96 \times 0.103$, which is statistically significant at the 5% level.

Before-After Studies

Accounting for change in traffic flow.

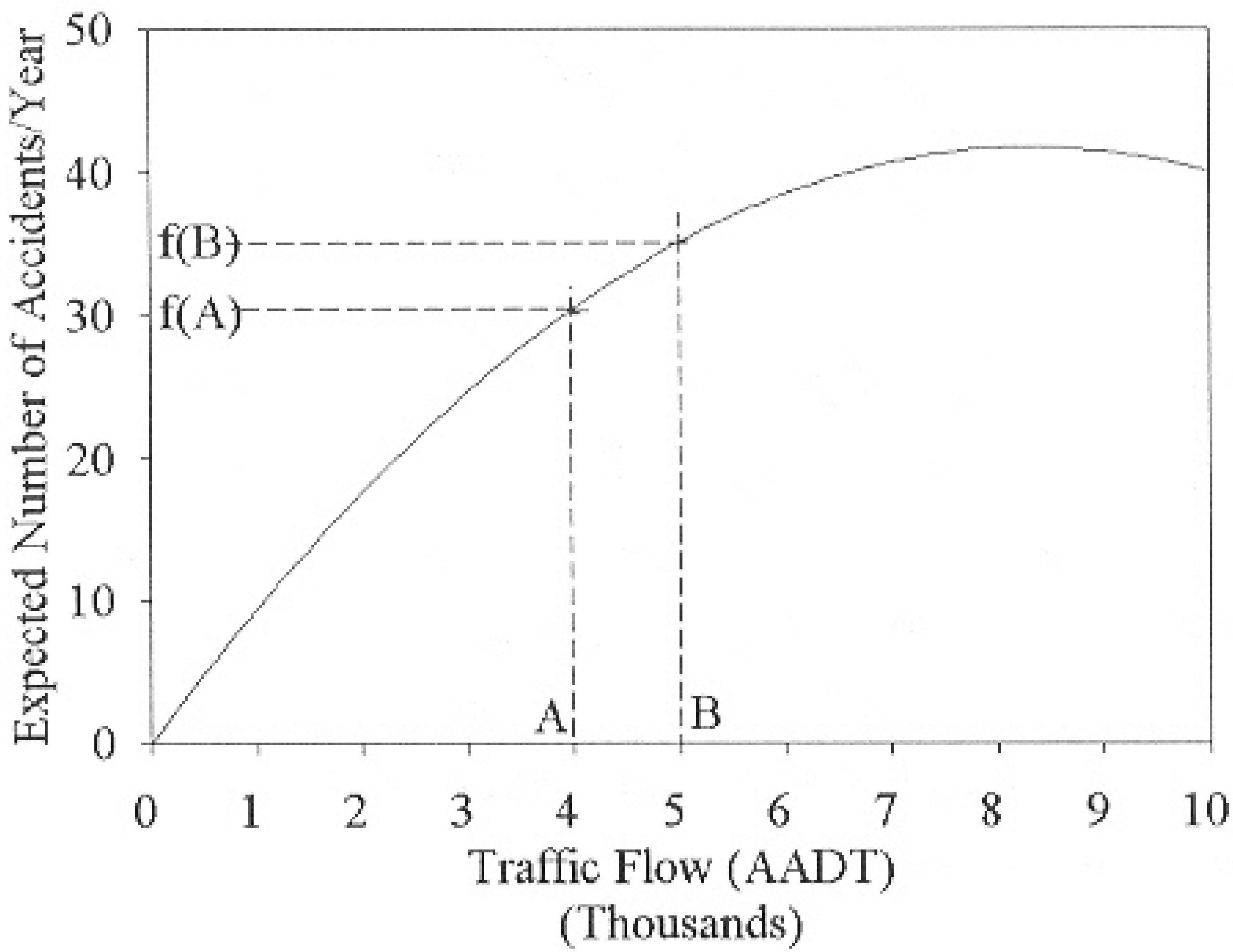
Both periods of same duration;
Traffic same in both periods; $\pi = \kappa$
Remaining factors same in both periods.

Traffic same in both periods;
Remaining factors same in both periods. $\pi = r_d \kappa$

Remaining factors same in both periods. $\pi = r_d \, r_{tf} \kappa$

Before-After Studies

Accounting for change in traffic flow.



Before-After Studies

Adjustment factor for change in traffic flow:

$$r_{tf} = f(A) \Big/ f(B)$$

$$\pi = r_d \times r_{tf} \times K$$

Note: $f\left(flow\right) = \beta_0 F^{\beta_1}$

Before-After Studies with Traffic Flow Factors

STEP 1 & STEP 2

Estimates of Coefficients

$$\lambda = \lambda$$

$$\pi = r_d r_{tf} \kappa$$

Estimates of Variances

$$Var\{\lambda\} = \lambda$$

$$Var\{\pi\} = r_d^2 \left[\hat{r}_{tf}^2 \times \kappa + \kappa^2 \times Var\{r_{tf}\} \right]$$

Described on next slide

Before-After Studies

Estimation of r_{tf}

$$r_{tf} = f(A_{avg}) \Big/ f(B_{avg})$$

$$Var\{r_{tf}\} \approx r_{tf}^2 \times \beta^2 \times \left[cv_{after}^2 + cv_{before}^2 \right]$$

Where *cv* is the percent of coefficient of variation (of the traffic flow). In practice, the percent coefficient of variation can be very difficult to obtain. Hence, if it is not available, values between 0.10 and 0.20 could be used in the equation above. It is recommended to conduct a sensitivity analysis to estimate how sensitive the *cv* is for different values.

Before-After Studies with Traffic Flow Factors

STEP 3 & STEP 4

$$\delta = \pi - \lambda$$

$$Var\{\delta\} = Var\{\pi\} + Var\{\lambda\}$$

$$\theta = \frac{\lambda}{\pi\left[1 + Var\{\pi\} / \pi^2\right]}$$

$$Var\{\theta\} \cong \frac{\theta^2\left[\left(Var\{\lambda\}/\lambda^2\right) + \left(Var\{\pi\}/\pi^2\right)\right]}{\left[1 + Var\{\pi\}/\pi^2\right]^2}$$

Before-After Studies

Example: Assume 572 vehicles were counted during a two-hour count for the before period and 637 were counted for the after period on a rural long-distance highway. Now assume that the functional relationship between crashes and flow is given by $f\left(flow\right)=\beta_0 F^{0.8}$.

Compute r_{tf} and $Var\{r_{tf}\}$, and assume the percent cv is 0.12 for both periods.

Before-After Studies

Coefficient of Variation

Example: Assume 572 vehicles were counted during a two-hour count for the before period and 637 were counted for the after period on a rural long-distance highway. Now assume that the functional relationship between crashes and flow is given by $f\left(flow\right)=\beta_0 F^{0.8}$.

Compute r_{tf} and $Var\{r_{tf}\}$, and assume the percent cv is 0.12 for both periods.

$$r_{tf}=\left({}^{637}\!/_{572}\right)^{0.8}=1.114^{0.8}=1.090$$

$$Var\{r_{tf}\}=1.09^2\times 0.8^2\times\left[0.12^2+0.12^2\right]=0.022$$

$$Var\{r_{tf}\}\approx r_{tf}^2\times\beta^2\times\left[cv_{after}^2+cv_{before}^2\right]$$

Before-After Studies

Continuing with the previous example. Now, assume that a road section has been resurfaced.

In the two-year 'before' period, 30 wet-pavement crashes were recorded on this section.

In the two-year 'after' period, 40 wet-pavement crashes were reported.

As before, 572 vehicles were counted during a two-hour count for the before period and 637 were counted for the after period. (F below)

The function relationship is still the same: $f\left(flow\right) = \beta_0 F^{0.8}$.

In addition, there were 50 wet-pavement days for the before period and 40 wet-pavement days for the after period.

Estimate δ, θ and the standard deviation of these estimates.

Before-After Studies

STEP 1: Estimate λ and π .

$$\lambda = 40$$

$$r_{tf} = 1.114^{0.8} = 1.090$$

$$r_d = 40/50 = 0.8$$

$$\pi = 0.8 \times 1.090 \times 30 = 26.16$$

Before-After Studies

STEP 2: Estimate $Var\{\lambda\}$ and $Var\{\pi\}$.

$$Var\{\lambda\} = 40$$

$$Var\{\pi\} = \left({}^{40}\!/_{50}\right)^2 \left[1.090^2 \times 30 + 30^2 \times 0.022\right]$$

$$Var\{\pi\} = 35.4$$

Before-After Studies

STEP 3: Estimate δ and θ .

$$\delta = 26.16 - 40 = -13.84$$

$$\theta = \frac{{}^{40}\!/_{26.16}}{\left[1 + {}^{35.4}\!/_{26.16^2}\right]}$$

$$\theta = 1.45$$

Before-After Studies

STEP 4: Estimate $Var\{\delta\}$ and $Var\{\theta\}$.

$$Var\{\delta\} = 35.4 + 40 = 75.4$$

$$Var\{\theta\} \approx \frac{1.45^2\left[\left(\frac{1}{40}\right)+\left(\frac{35.4}{26.16^2}\right)\right]}{\left[1+\frac{35.4}{26.16^2}\right]^2}$$

$$Var\{\theta\} \approx 0.144$$

Before-After Studies with Comparison Group

Comparison groups are used to capture changes that change over time (described previously). The two main assumptions are:

a) The sundry factors that affect safety have changed from the before to the after period in the same manner on both treatment and the comparison group

b) This change in the sundry factors affects the treatment and the comparison group in the same way.

r_c = the comparison ratio; the ratio of the expected number of "after" to the expected number of "before" target crashes of the comparison group

Before-After Studies with Comparison Group

Crash Counts and Expected Values

	Treatment Group	Comparison Group
Before	κ	μ
After	λ	ν

Before-After Studies with Comparison Group

Let us define the following notations:

$$r_c = \nu / \mu$$

The ratio of the expected crash counts for the comparison group

$$r_t = \pi / \kappa$$

The ratio of the expected crash counts for the treatment group

The hope is that $r_t = r_c$ $\quad \therefore \pi = r_c \kappa = r_t \kappa$

$$\omega = r_c / r_t$$

Odd's ratio

Time periods need to be the same for both the comparison and treatment groups

Before-After Studies with Comparison Group

STEP 1 & STEP 2

Estimates of Coefficients

$$\lambda = \lambda$$

$$r_t = r_c = (\nu / \mu) / (1 + 1/\nu)$$

$$r_t \approx \nu / \mu$$

$$\pi = r_t \kappa$$

Estimates of Variances

$$Var\{\lambda\} = \lambda$$

$$Var\{r_t\} / r_t^2 \approx 1/\mu + 1/\nu + Var\{\omega\}$$

Using 0.001 is good for most cases.

$$Var\{\pi\} = \pi^2 \left[1/\kappa + Var\{r_t\} / r_t^2 \right]$$

Before-After Studies with Comparison Group

STEP 3 & STEP 4

$$\delta = \pi - \lambda$$

$$Var\{\delta\} = Var\{\pi\} + Var\{\lambda\}$$

$$\theta = \frac{\lambda}{\pi\left[1 + Var\{\pi\} / \pi^2\right]}$$

$$Var\{\theta\} \cong \frac{\theta^2\left[\left(Var\{\lambda\}\middle/\lambda^2\right) + \left(Var\{\pi\}\middle/\pi^2\right)\right]}{\left[1 + Var\{\pi\}\middle/\pi^2\right]^2}$$

Before-After Studies with Comparison Group

Exercise 7.3

Redo Exercise 7.1, but consider the 25 sites that were included in the comparison group. Use 5 years before and 3 years after.

(Note: the value of the following parameters shows the summation over 15 sites.)

First, calculate r_c

$$r_c = \frac{\nu}{\mu}$$

$$r_c = \frac{259}{405}$$

$$r_c = 0.64$$

Estimate π, λ, $Var(\lambda)$

Site ID	r_c	π	λ	$Var(\lambda)$
1	0.64	7.7	5	5
2	0.64	9.6	9	9
3	0.64	10.2	5	5
4	0.64	10.2	5	5
5	0.64	16.6	9	9
6	0.64	9.0	5	5
7	0.64	16.0	12	12
8	0.64	12.2	9	9
9	0.64	12.2	16	16
10	0.64	11.5	14	14
11	0.64	18.5	8	8
12	0.64	16.6	12	12
13	0.64	3.8	11	11
14	0.64	9.0	8	8
15	0.64	19.8	12	12
Sum		182.9	140	140

Before-After Studies with Comparison Group

Calculate $Var(\pi)$

$$Var(r_t)/r_t^2 \approx 1/\mu + 1/\nu + Var(\omega)$$

$$Var(r_t)/r_t^2 \approx 1/405 + 1/259 + 0.001$$

$$Var(r_t)/r_t^2 \approx 0.0073$$

$$Var(\pi) = \pi^2 \left[1/\kappa + Var(r_t)/r_t^2\right]$$

$$Var(\pi) = 182.9^2 [1/286 + 0.0073]$$

$$Var(\pi) = 362.2$$

Calculate δ

$$\delta = 42.9$$

Before-After Studies with Comparison Group

There is a reduction of 42.9 in the expected number of crashes.

Calculate θ (adjust for a sample size below 500)

$$\theta = 0.77$$

There is a reduction of 23% in the expected number of crashes. This value is larger than the reduction observed in Exercise 7.1. The data from the comparison group shows a small reduction in the number of crashes as well, which explains why θ decreased from 82% to 77%.

Calculate $Var(\delta)$ and $SD(\delta)$

$$Var(\delta) = 502.2$$
$$SD(\delta) = 22.4$$

The reduction is $42.9 \pm 1.96 \times 22.4$, which is not statistically significant at the 5% level (almost at the boundary).

Calculate $Var(\theta)$ and $SD(\theta)$

$$Var(\theta) \cong 0.010$$
$$SD(\theta) = 0.101$$

The reduction is $0.77 \pm 1.96 \times 0.101$, which is statistically significant at the 5% level.