Introduction to road safety: Basic concepts, data and statistical analysis

Luis Miranda-Moreno, PhD
Assistant Professor
Department of Civil Engineering
E-mail: luis.miranda-moreno@mcgill.ca
Outline

1. Basic concepts
2. Data preparation and introduction to road safety methods
1. basic definitions
Factors influencing crash occurrence

Road crashes are multi-factorial events

- **Driver behavior**
  - Speed,
  - Alcohol & drug use
  - Seatbelt and helmet use
  - Fatigue

- **Vehicle conditions**
  - Anti-lock brake
  - Vehicle devices
  - Maintenance

- **Roadway conditions**:
  - Vertical & horizontal alignment
  - Surface conditions
  - Signage, traffic controls, etc.

- **Traffic exposure**
  - Interactions/volumes

- **Weather conditions**
  - Visibility,
  - Precipitation,
  - Winter
3Es approach

In the “triple E” system, each “E” stands for:
1. Engineering (road engineering and vehicle engineering),
2. Education (training, traffic education) and
3. Enforcement.

Road Safety Manual, PIARC, 2011
General definitions

- **Crashes (crashes, collisions):** outcomes of interactions and sequences of actions between road users and road environment.

  \[ \text{Interactions } \Rightarrow \text{dangerous situations } \Rightarrow \text{crashes} \]
General definitions – cont.’

- **Crash occurrence** (frequency): Number of crashes occurring at a particulate site, facility or network in a period of time (this can be in one-year period or several years).

- **Crash injury severity**: Is the level of injury or property damage. Typical injury scale from police reports:
  - fatal,
  - major injuries,
  - minor injuries,
  - Property damage only (PDO).

As transportation specialist, we want to target both: crash occurrence and consequences (total risk)
From undisturbed passages to crashes

Hyden’s classic pyramid-model (1985)

Typical crash classification:

- Crashes
- Serious conflicts
- Slight (minor) conflicts
- Potential conflicts
- Undisturbed passages

- Fatal
- Severe injury
- Minor injury
- Damage only
Alternative methods

• Crash-based approach:
  + Well recognized in the literature and practice
  - “Crashes need to happen to be recorded”

• Surrogate approach (speed, conflicts):
  This can be seen as a complementary approach. It assumes existence of a casual link to expected crash frequency
  - Needs more validation
  + “we do not need to wait for crashes to happen”
Traffic exposure and crash occurrence

Aggregate indicators:
Number of passages (traffic volume) in a given facility (e.g., intersection, midblock location) during a given period of time.

*Exposure* = *f*(number of vehicles, pedestrians and bikes)
e.g., Typical exposure indicator: Average Annual Daily Traffic AADT)

Mean number of crashes = flow^α

Rate = crashes/ flow^α

(PIARC, 2004)
Literature

- **Aggregate indicators of exposure:**
  Previous research has focused on the relationship between vehicle-bicycle crash frequency with traffic and bike volumes (Elvik, 2009, Strauss & Miranda-Moreno, 2012).

\[ \lambda_i = \alpha_0 Z_i^{\alpha_1} F_i^{\alpha_2} \]

- \( \lambda_i \) = mean number of vehicle-cyclist collisions at a given site \( i \)
- \( Z_i \) = Bike flow at \( i \)
- \( F_i \) = Traffic volume at \( i \)
Non-linear association (Elvik, 2009)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Country</th>
<th>Study units</th>
<th>Sample size</th>
<th>Number of accidents</th>
<th>Type of accident</th>
<th>Measure of exposure</th>
<th>Exponent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bråde, Larsson</td>
<td>1993</td>
<td>Sweden</td>
<td>Junctions</td>
<td>285</td>
<td>165</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, crossing pedestrians</td>
<td>0.50</td>
</tr>
<tr>
<td>Bråde, Larsson</td>
<td>1993</td>
<td>Sweden</td>
<td>Junctions</td>
<td>377</td>
<td>432</td>
<td>Cyclist</td>
<td>Entering motor vehicles, entering cyclists</td>
<td>0.52</td>
</tr>
<tr>
<td>Ledin, Gärder, Pulikkinen</td>
<td>1998</td>
<td>Sweden</td>
<td>Junctions</td>
<td>749</td>
<td>276</td>
<td>Cyclist</td>
<td>Entering motor vehicles, crossing pedestrians</td>
<td>0.66</td>
</tr>
<tr>
<td>Ledin</td>
<td>2002</td>
<td>Canada</td>
<td>Junctions</td>
<td>126</td>
<td>27</td>
<td>Pedestrian</td>
<td>Left turning motor vehicles, crossing pedestrians</td>
<td>1.19</td>
</tr>
<tr>
<td>Lyon, Persaud</td>
<td>2002</td>
<td>Canada</td>
<td>Junctions</td>
<td>684</td>
<td>5280</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.57</td>
</tr>
<tr>
<td>Lyon, Persaud</td>
<td>2002</td>
<td>Canada</td>
<td>Junctions</td>
<td>263</td>
<td>1065</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.40</td>
</tr>
<tr>
<td>Lyon, Persaud</td>
<td>2002</td>
<td>Canada</td>
<td>Junctions</td>
<td>122</td>
<td>159</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.53</td>
</tr>
<tr>
<td>Lyon, Persaud</td>
<td>2002</td>
<td>Canada</td>
<td>Junctions</td>
<td>123</td>
<td>319</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.58</td>
</tr>
<tr>
<td>Jacobsen, Jacobsen</td>
<td>2003</td>
<td>United States</td>
<td>Cities</td>
<td>68</td>
<td></td>
<td>Pedestrian</td>
<td>Share of working trips on foot bicycle</td>
<td>0.41</td>
</tr>
<tr>
<td>Jacobsen, Jacobsen</td>
<td>2003</td>
<td>United States</td>
<td>Cities</td>
<td>68</td>
<td></td>
<td>Cyclist</td>
<td>Share of working trips on foot bicycle</td>
<td>0.31</td>
</tr>
<tr>
<td>Jacobsen</td>
<td>2003</td>
<td>Denmark</td>
<td>Towns</td>
<td>47</td>
<td></td>
<td>Pedestrian</td>
<td>Kilometres walked per inhabitant per day</td>
<td>0.36</td>
</tr>
<tr>
<td>Jacobsen</td>
<td>2003</td>
<td>Denmark</td>
<td>Towns</td>
<td>47</td>
<td></td>
<td>Cyclist</td>
<td>Kilometres cycled per inhabitant per day</td>
<td>0.44</td>
</tr>
<tr>
<td>Jacobsen</td>
<td>2003</td>
<td>14 European</td>
<td>Country</td>
<td>14</td>
<td></td>
<td>Cyclist</td>
<td>Kilometres cycled per inhabitant per day</td>
<td>0.58</td>
</tr>
<tr>
<td>Jacobsen</td>
<td>2003</td>
<td>8 European</td>
<td>Country</td>
<td>8</td>
<td></td>
<td>Pedestrian</td>
<td>Trips on foot per inhabitant per day</td>
<td>0.13</td>
</tr>
<tr>
<td>Jacobsen</td>
<td>2003</td>
<td>8 European</td>
<td>Country</td>
<td>8</td>
<td></td>
<td>Cyclist</td>
<td>Trips on bicycle per inhabitant per day</td>
<td>0.48</td>
</tr>
<tr>
<td>Robinson</td>
<td>2005</td>
<td>Australia</td>
<td>States</td>
<td>7</td>
<td></td>
<td>Cyclist</td>
<td>Kilometres cycled per inhabitant per day</td>
<td>0.52</td>
</tr>
<tr>
<td>Jonsson</td>
<td>2005</td>
<td>Sweden</td>
<td>Road sections</td>
<td>393</td>
<td>130</td>
<td>Pedestrian</td>
<td>Motor vehicle kilometres, pedestrians crossing and walking along road</td>
<td>0.83</td>
</tr>
<tr>
<td>Jonsson</td>
<td>2005</td>
<td>Sweden</td>
<td>Road sections</td>
<td>393</td>
<td>343</td>
<td>Cyclist</td>
<td>Motor vehicle kilometres, cyclists crossing and riding along road</td>
<td>0.76</td>
</tr>
<tr>
<td>Geyer et al.</td>
<td>2006</td>
<td>Oakland</td>
<td>Junctions</td>
<td>247</td>
<td>185</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, crossing pedestrians</td>
<td>0.16</td>
</tr>
<tr>
<td>Harwood et al.</td>
<td>2008</td>
<td>United States</td>
<td>Junctions</td>
<td>450</td>
<td>728</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.05</td>
</tr>
<tr>
<td>Harwood et al.</td>
<td>2008</td>
<td>United States</td>
<td>Junctions</td>
<td>1433</td>
<td>4824</td>
<td>Pedestrian</td>
<td>Entering motor vehicles, entering pedestrians</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Mean (simple) 0.57 0.50 0.48
Disaggregate exposure indicators
Number of interactions per movements & phase

Miranda-Moreno, Strauss (2012)
Disaggregate exposure indicator

Number of red-light crossings (violations)

Miranda-Moreno, Strauss (2012)
Other exposure indicators

Based on the crash risk of crossing (*Routledge et al. 1976; Lassarre et al. 2007)*:

- Average speed of flow
- Average traffic gap
- Average crossing time (length of crossing)
- Vehicle length

“Proportion of space that is NOT available to the pedestrian for crossing the road freely and safely”.
Literature

Factors positively or negatively associated to the crash occurrence in non-motorized traffic at intersections:

- Exposure: Volume intensity, turning movements (in particular right turns)
- Road width / number of lanes / parking
- Bus stop presence / parking entrances, etc.
- Pedestrian signals
- Etc.

Engineering treatments

Exposure:

- Separating / eliminating conflicting movements:
  - Exclusive or half phase for pedestrians/bikes
  - Restricting vehicle right-turns

- Integrating ped/bike delays into traffic control design
- Change intersection setting (e.g., number of approaches, roundabouts)
Engineering treatments – cont’

Traversed distance/visibility/turning speed

✓ Add curb extensions (speed reduction, visibility)
✓ Eliminate parking in the proximity
✓ Add two-stage turn queue boxes
✓ Bike/pedestrian green light, etc.

See: NACTO Urban Bikeway Design Guide, Vélo Québec guidelines, Dutch design manual
Engineering treatments – cont’

Bike boxes

Corner Island

Cycle Track Bend Away from Street

Pedestrian platform with signal

http://wiki.coe.neu.edu/groups/nl2011transpo/wiki/bd54f/Bicycling_Facilities_in_Holland.html
The traffic “challenge”

*Vehicular Capacity* (\(c\)): Traffic volume (veh/h) that can pass through an intersection from a lane or group of lanes:

\[ c = s \times \frac{g}{C} \]

Where:
- \(c\) = max. hourly volume that can pass through an intersection from a lane or group of lanes, in veh/h
- \(s\) = saturation flow rate in veh/h (\(s = 3600/\text{sat headway}\))
- \(g/C\) = ratio of effective green time to cycle length.

Reducing vehicle space and/or green time will affect “\(s\)” and “\(g\)”

This can increase delays and deteriorate level of service for drivers

*Fred Mannering, et al. 2010*
2. Data preparation and statistical analysis
Transportation management process

Conceptual Approaches to Road Safety
(Highway Safety Manual, AASHTO, 2010).
Data for safety analysis

Traffic monitoring tools

Manual data collection

Weather data

Expansion factors and manual counts

Monitoring and safety analysis

Intersection inventory

Crash data
Spatial traffic and crash data

Pedestrian ADDT

Vehicular ADDT

Injuries per intersection
Traffic safety studies

Studies can be divided in two:
- Cross-sectional studies / control-case study
- Before-after observational studies

Applications:
- Safety effectiveness: before-after studies
- Network screening: mapping crash risk

Methods for crash risk estimation:
- Raw risk approach
- Model-based approach
“Raw” crash-based method

\[ R_i = \frac{y_i \times 10^6}{365(T_i \times AADT_i)} \]

\( R_i \) = injury rate of intersection i (injuries per million of cyclists per unit of time)

\( y_i \) = bicycle injury frequency at intersection i during \( T_i \)

\( T_i \) = period of analysis (years)

\( AADT_i \) = Average annual daily bicycle volume of intersection i

*Road Safety Manual, PIARC, 2011*
Example 1: Control-case study on bike lanes

Evaluate the cyclist risk (injury rates) in the following 3 street sections with bicycle lanes. Compare the risk with the similar 3 street sections without bicycle lanes. Determine the safety effectiveness of bike lanes.

<table>
<thead>
<tr>
<th>ID</th>
<th>Section</th>
<th>Years</th>
<th>L. (km)</th>
<th>AADT</th>
<th>Number of injuries</th>
</tr>
</thead>
<tbody>
<tr>
<td>Facilities</td>
<td>1 Milton</td>
<td>1.5</td>
<td>0.52</td>
<td>1,797</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>2 Hutchison (south)</td>
<td>2.5</td>
<td>0.78</td>
<td>2,343</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>3 Prince Arthur</td>
<td>3.5</td>
<td>0.96</td>
<td>805</td>
<td>3.0</td>
</tr>
<tr>
<td>Controls</td>
<td>1 Control 1(Hutchison north)</td>
<td>2.5</td>
<td>0.75</td>
<td>1,958</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>2 Control 2 (Brebeuf)</td>
<td>3.5</td>
<td>1.1</td>
<td>1801</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>3 Control 3 (Boyer)</td>
<td>3.5</td>
<td>1.5</td>
<td>803</td>
<td>7.0</td>
</tr>
</tbody>
</table>
Solution

\[ R_i = \frac{y_i \times 10^6}{365(T_i \times L_i \times AADT_i)} \]

Injuries per million of bike-km

<table>
<thead>
<tr>
<th>ID</th>
<th>Section</th>
<th>Years</th>
<th>L. (km)</th>
<th>AADT</th>
<th>Number of injuries</th>
<th>Injury Rates</th>
<th>Average values</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Milton</td>
<td>1.5</td>
<td>0.52</td>
<td>1797</td>
<td>2.0</td>
<td>3.91</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Hutchison (south)</td>
<td>2.5</td>
<td>0.78</td>
<td>2343</td>
<td>2.0</td>
<td>1.20</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Prince Arthur</td>
<td>3.5</td>
<td>0.96</td>
<td>805</td>
<td>3.0</td>
<td>3.04</td>
<td>2.72</td>
</tr>
<tr>
<td></td>
<td>Control 1 (Hutchison north)</td>
<td>2.5</td>
<td>0.75</td>
<td>1958</td>
<td>3.0</td>
<td>2.24</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Control 2 (Brebeuf)</td>
<td>3.5</td>
<td>1.1</td>
<td>1801</td>
<td>7.0</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Control 3 (Boyer)</td>
<td>3.5</td>
<td>1.5</td>
<td>803</td>
<td>7.0</td>
<td>4.55</td>
<td>3.18</td>
</tr>
</tbody>
</table>

Average reduction = \([1 - (3.18/2.72)]*100 = - 17\%\)
Elements to take into account:

• **Average Annual Daily Traffic (ADDT):** How to determine this using short-term counts considering temporal/weather effect?

• **Crash (injuries) data:** How many years of data? Low-mean problem?

• **Temporal changes on AADT and y_i:** E.g., before and after the installation of a bike infrastructure.
Adjustment factors for different traffic patterns

<table>
<thead>
<tr>
<th>Type</th>
<th>Hourly Profiles</th>
<th></th>
<th>Daily Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Weekday</td>
<td>Weekend</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.15</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>0.10</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.15</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>0.10</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>0.15</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>0.10</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.05</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Miranda-Moreno, Nosal, 2012
Volume trends in bicycle facilities in Montreal

Average Seasonal Hourly Traffic

Average hourly volume

Hours

- 2012
- 2011
- 2010
- 2009
- 2008
Before-after traffic volumes

1) Bicycle lanes, Ave. Laurier, Montreal
2) Cycle track, Ave. Laurier, Ottawa

Increase: > 100% with respect to May 2011
Increase: ~ 50% with respect to May 2011
Example 2: Bicycle traffic

To improve the traffic safety of cyclists, the city of Ottawa recently installed a bicycle facility on Ave Laurier – a bidirectional cycle track. The following data was collected before and after the installation in 2011:

- April 29 (Fri): 150 bikes from 8:00 to 10:00am
- May 11 (Wed): 95 bikes from 9:00 to 10:00am
- May 16 (Mon): 211 bikes from 8:00 to 10:00am
- June 15 (Wed): 284 bikes from 8:00 to 10:00am
- June 30 (Fri): 125 bikes from 9:00 to 10:00am
- July 25 (Mon): 299 from 8:00 to 10:00

Weather conditions: no rain, temp between 15°C and 25 °C

Question: The facility started operating on June 06, 2011. Determine the increase or decrease in terms of AADT according to the manual counts (adjusting only for temporal trends).

Use the expansion factors obtained from a counting stations located in a bike facility with similar traffic patterns in the same city.
Expansion factors from automatic counts

Hourly traffic variations (in percentage out of 100)

<table>
<thead>
<tr>
<th>Hour</th>
<th>Weekday</th>
<th>Saturday</th>
<th>Sunday</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1.03</td>
<td>3.08</td>
<td>3.26</td>
</tr>
<tr>
<td>1</td>
<td>0.48</td>
<td>1.84</td>
<td>2.07</td>
</tr>
<tr>
<td>2</td>
<td>0.24</td>
<td>1.10</td>
<td>1.26</td>
</tr>
<tr>
<td>3</td>
<td>0.20</td>
<td>0.93</td>
<td>1.12</td>
</tr>
<tr>
<td>4</td>
<td>0.18</td>
<td>0.43</td>
<td>0.48</td>
</tr>
<tr>
<td>5</td>
<td>0.58</td>
<td>0.40</td>
<td>0.39</td>
</tr>
<tr>
<td>6</td>
<td>1.91</td>
<td>0.68</td>
<td>0.64</td>
</tr>
<tr>
<td>7</td>
<td>5.72</td>
<td>1.26</td>
<td>1.14</td>
</tr>
<tr>
<td>8</td>
<td>11.24</td>
<td>2.46</td>
<td>2.10</td>
</tr>
<tr>
<td>9</td>
<td>7.07</td>
<td>3.66</td>
<td>3.31</td>
</tr>
<tr>
<td>10</td>
<td>4.18</td>
<td>4.64</td>
<td>4.55</td>
</tr>
<tr>
<td>11</td>
<td>4.22</td>
<td>5.80</td>
<td>6.03</td>
</tr>
<tr>
<td>12</td>
<td>4.77</td>
<td>6.75</td>
<td>7.27</td>
</tr>
<tr>
<td>13</td>
<td>5.11</td>
<td>7.71</td>
<td>8.74</td>
</tr>
<tr>
<td>14</td>
<td>5.11</td>
<td>8.49</td>
<td>9.56</td>
</tr>
<tr>
<td>15</td>
<td>5.80</td>
<td>8.49</td>
<td>9.41</td>
</tr>
<tr>
<td>16</td>
<td>8.01</td>
<td>8.15</td>
<td>8.58</td>
</tr>
<tr>
<td>17</td>
<td>10.71</td>
<td>7.40</td>
<td>7.56</td>
</tr>
<tr>
<td>18</td>
<td>7.23</td>
<td>5.89</td>
<td>5.84</td>
</tr>
<tr>
<td>19</td>
<td>4.78</td>
<td>4.96</td>
<td>4.42</td>
</tr>
<tr>
<td>20</td>
<td>3.75</td>
<td>4.08</td>
<td>3.81</td>
</tr>
<tr>
<td>21</td>
<td>3.35</td>
<td>4.13</td>
<td>3.30</td>
</tr>
<tr>
<td>22</td>
<td>2.44</td>
<td>3.88</td>
<td>2.79</td>
</tr>
<tr>
<td>23</td>
<td>1.89</td>
<td>3.79</td>
<td>2.39</td>
</tr>
</tbody>
</table>

Day | Daily expansion factors |
--- | ------------------------|
Sunday | 0.68 |
Monday | 1.03 |
Tuesday | 1.16 |
Wednesday | 1.19 |
Thursday | 1.12 |
Friday | 1.08 |
Saturday | 0.74 |

Month | Monthly expansion factors |
--- | --------------------------|
April | 0.80 |
May | 1.09 |
June | 1.17 |
July | 1.27 |
August | 1.20 |
September | 1.13 |
October | 0.67 |
November | 0.55 |

Winter months factors are 0
Solution

\[ \text{ADDT} = V \times [(1/F_h) \times (1/F_D) \times (1/F_M)] \]

*Fh, F_D, F_M - hourly, daily and monthly expansion factors*

<table>
<thead>
<tr>
<th>Hour</th>
<th>Day / month</th>
<th>Volume</th>
<th>Hourly factors</th>
<th>Daily factors</th>
<th>Monthly factors</th>
<th>Daily average (corrected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8 - 10</td>
<td>Fri / April</td>
<td>150</td>
<td>0.1831</td>
<td>1.08</td>
<td>0.8</td>
<td>948</td>
</tr>
<tr>
<td>9 - 10</td>
<td>Wed / May</td>
<td>95</td>
<td>0.0707</td>
<td>1.19</td>
<td>1.09</td>
<td>1036</td>
</tr>
<tr>
<td>8 - 10</td>
<td>Mon / May</td>
<td>211</td>
<td>0.1831</td>
<td>1.03</td>
<td>1.09</td>
<td>1026</td>
</tr>
</tbody>
</table>

Average daily (before): 1004

<table>
<thead>
<tr>
<th>Hour</th>
<th>Day / month</th>
<th>Volume</th>
<th>Hourly factors</th>
<th>Daily factors</th>
<th>Monthly factors</th>
<th>Daily average (corrected)</th>
</tr>
</thead>
<tbody>
<tr>
<td>After</td>
<td>8-10</td>
<td>Wed / June</td>
<td>284</td>
<td>0.1831</td>
<td>1.19</td>
<td>1.17</td>
</tr>
<tr>
<td>9 - 10</td>
<td>Fri / June</td>
<td>125</td>
<td>0.0707</td>
<td>1.08</td>
<td>1.17</td>
<td>1399</td>
</tr>
<tr>
<td>8 - 10</td>
<td>Mon / July</td>
<td>299</td>
<td>0.1831</td>
<td>1.03</td>
<td>1.27</td>
<td>1248</td>
</tr>
</tbody>
</table>

Average daily (after): 1254

Estimated % increase: 0.25
Shortcoming of “raw crash” risk indicator

a) Uncertainty is not taken into account

b) They do not consider the possibility of a **non-linear relationship** between crashes and traffic exposure

c) Variations in **roadway characteristics**: raw estimates ignore the effect of site-specific attributes (e.g., geometry and signalization factors)

So… it can lead to wrong estimations

*Road Safety Manual, PIARC, 2011*
2.1 Statistical models for crash frequency analysis
Poisson Approximation

Since a (crash) event has a very low probability of occurrence and a large number of trials exist (e.g. million entering vehicles, vehicle-miles-traveled, etc.), the binomial distribution is approximated by a Poisson distribution.

\[
P(Y = n) = \left(\frac{N}{n}\right)\left(\frac{\mu}{N}\right)^n\left(1 - \frac{\mu}{N}\right)^{N-n} \approx \frac{\mu^n}{n!} e^{-\mu}
\]

Where,

\( n = 0, 1, 2, \ldots, N \) (the number of successes or crashes)

\( \mu = \) the mean of a Poisson distribution
Negative Binominal (Poisson-Gamma) model

\[ Y_i | \theta_i \sim \text{Poisson} (T_i \theta_i) \]
\[ \sim \text{Poisson} (T_i \mu_i e^{\varepsilon_i}) \]

Random effect
\[ e^{\varepsilon_i} \sim \text{Gamma}(\phi, \phi) \]

where: \( Y_i = \) number of crashes at site \( i \)
\( T_i = \) observation time at \( i \)
\( \theta_i = \) mean number of crashes at \( i \)
\( \mu_i = \beta_0 F_{1i}^{\beta_1} F_{2i}^{\beta_2} \exp(\beta_3 x_3 + ... + \beta_k x_k) \)
\( F_{1i}, F_{2i} = \) Traffic flows,
\( \beta_0, \beta_1, ..., \beta_k = \) regression parameters
\( x_k = \) geometry/traffic control attributes
Typical safety performance functions (SPF)

\[ \mu_i = F_{1i}^{\beta_1} F_{2i}^{\beta_2} \exp(\beta_0 + \beta_3 X_{3i} + \ldots + \beta_k X_{ki}) \]  

(for intersections)

\[ \mu_i = L_i (F_{1i} + F_{2i})^{\beta_1} \exp(\beta_0 + \beta_2 x_{i2} + \ldots + \beta_k x_{ik}) \]  

(for sections)

(PIARC, 2004)
Posterior analysis using NB model

The Poisson/Gamma or Negative Binomial model can be written as (considering $T$ constant):

$$ Y_i \mid \theta_i \sim \text{Poisson}(\theta_i) $$

$$ \theta_i \sim \text{Gamma}(\phi, \phi/\mu_i) $$

Applying Bayes’ theorem:

$$ p(\theta_i \mid y_i) \propto f(y_i \mid \theta_i) \pi(\theta_i) $$

$$ \propto \frac{e^{-\theta_i} (\theta_i)^{y_i}}{y!} \left(\frac{\phi}{\mu_i}\right)^\phi \theta_i^{\phi-1} e^{-(\phi/\mu_i)\theta_i} \frac{\Gamma(\phi)}{\Gamma(\phi)} $$

or $$ \theta_i \mid y_i \sim \text{Gamma}(y_i + \phi, 1 + \phi/\mu_i) $$
EB estimator based on NB model

Knowing that \( p(\theta_i | y_i) \) is a Gamma distribution with shape \( a = (y_i + \phi) \) and scale parameter \( b = (1 + \phi/\mu_i) \), the posterior mean and variance of \( \theta_i \) is given by:

\[
E(\theta_i \mid y_i) = \frac{y_i + \phi}{1 + \phi/\mu_i} \quad \text{and} \quad \text{Var}(\theta_i \mid y_i) = \frac{y_i + \phi}{(1 + \phi/\mu_i)^2}
\]

\[
\text{EB}_\theta = E(\theta_i \mid y_i) = (1 - w_i)y_i + w_i\mu_i
\]

with \( w_i = \phi/(\phi + \mu_i) \)

This is known in the safety literature as the EB estimator and has been widely implemented by researchers and government agencies for safety analysis.
Example 3: bicycle injury analysis at signalized intersections

Carry on the following tasks using a sample of signaled intersections.
- Define risk exposure measures according to the definitions provided before.
- Develop collision occurrence models and select the most appropriate ones (report only the best models).
- Identify the main contributing factors (traffic conditions, geometric, controls, etc.) using the parameters, confidence intervals, elasticities.

Data description:

Collision data

- Bike injuries: represents the traffic-related bicycle injuries for a six-year period and includes all vehicle-bicycle collisions within a 15-meter radius.
- AADT_bike: Average annual daily traffic for bikes
- AADT_vehicle: Average annual daily traffic for motor vehicles.
- AADTs were obtained using 8-hour manual counts. They have been expanded using appropriate factors (temporal and weather factors for bikes).
- Geometry variables
2.2 Before-After Studies
Before-after studies

• Types of methods
  – Naïve before-after method
  – Before-after method with control group
  – Empirical Bayes approach (with control group)

• Important Issues
  – Selection bias
  – Regression-to-the-mean
  – Non-linear relationship between crashes and traffic flow
  – Low mean
Site Selection Bias

Sites that are treated are not randomly selected. Therefore, treatment effectiveness cannot be estimated directly without taking this into account.

Safety inspections and potential improvement
Regression-to-the-mean

- This consists of the general tendency of **extreme values** to regress to mean values.

- Then, if one treats hazardous locations with high observed or estimated crash history, the effect **RTM** should be considered.

![Figure 11.1. How accident counts regress to the mean.](image)

*Hauer, 1997*
Before-after methods for evaluation of treatment effects

1. Naïve before-after approach
2. Before-after approach using a control group
3. Empirical Bayes approach using a control group
The change in crashes for a treated intersection is given by:

\[ \text{Treatment effect (δ)} = K - \lambda \]

where:

- \( K \) = # of crashes in the before period
- \( \lambda \) = # of crashes in the after period

Assuming the observation period (time) before and after the treatment is the same.
Before-after study using a control group

\[ \pi = \text{Estimated no. of crashes without treatment} \]

Reduction of crashes due to treatment implementation

\[ \pi \] is estimated using a site or a group of similar sites in which non-treatment have been applied

\[ \text{Treatment effect } (\delta) = \pi - \lambda \]

Challenge: how to select an appropriate control group
Before-after study using a control group

$$\pi = r_t \times K$$

$r_t =$ ratio between the no. of crashes in the after / no. of crashes in the before period on the comparison site (or group)

$\pi =$ Estimated no. of crashes without treatment

Assumption: factors affecting safety changed in the same way on both treatment and comparison group
Example 4: Before-After study using a control group

A “Bike box” was installed to increase visibility of cyclists and reduce conflicts between motor vehicles and cyclists, particularly in potential “right-hook” situations. Injuries in the treated and control intersections have been recorded in the 3 years before and after the installation:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intersection</td>
<td>AADT</td>
</tr>
<tr>
<td>Treated site</td>
<td></td>
<td>925</td>
</tr>
<tr>
<td>Control site</td>
<td></td>
<td>825</td>
</tr>
</tbody>
</table>
Solution:

Calculating injury rates

<table>
<thead>
<tr>
<th></th>
<th>K</th>
<th>( \lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated intersection</td>
<td>5.1</td>
<td>3.6</td>
</tr>
<tr>
<td>Control intersection</td>
<td>5.2</td>
<td>4.2</td>
</tr>
</tbody>
</table>

\[ r_t = \text{Injury rate in the after /injury rate in the before period (comparison site)} = \frac{4.2}{5.2} = 0.8 \]

\[ \pi = K \times r_t = 5.1 \times 0.8 = 4.1 \]

Estimated crash reduction = 4.1 – 3.6 = 0.5 injuries
Before-after study using EB approach

\[ \pi = EB \times r_t \]

\( r_t = \) ratio between the no. of crashes in the after / no. of crashes in the before period on the comparison site (or group)

\( EB = \) Expected number of “before” crashes on the treated site (or group) using the posterior distribution of crashes

Reduction of crashes due to treatment implementation
Final remarks

- Keep in mind the quality of data:
  - Not all reportable crashes are actually reported
  - Crash location can be inaccurate
  - Bike/ped demand is very sensitive to land-use, weather, etc.

- In non-motorized safety: low mean can be an issue

- Surrogate analysis could be a good complement: more validation is needed.

- A lot of opportunities for research