Statistical Analysis of Vehicle Crash Incidents at the Minnesota USA Interstate 494 and Interstate 35 West Interchange using Chi-Square Statistic Goodness-of-Fit Testing

Candace Roberts
Department of Resource Analysis, Saint Mary's University of Minnesota, Minneapolis, MN 55404

Keywords: Vehicle Crash Incidents, Minneapolis I-494 and I-35W Cloverleaf Interchange, Contributing Factors, Chi-Square Goodness-of-Fit, Contingency Tables

Abstract

The Minneapolis, Minnesota (USA) Interstate 494 and Interstate 35 West cloverleaf interchange experiences a continued overcapacity daily traffic volume, creating the lowest safety ranking of highway interchanges in the metropolitan area. During a 10-year period from 2006 through 2015, an average 8.45 monthly crash incidents were reported. The majority of crashes occurred during good weather conditions and were attributed to driver behavior factors. This study utilized contingency tables and Chi-Square Goodness-of-Fit tests to determine associations and statistical equality of contributing factors that influenced crash incidents. Significant findings point to contributing factors for which improvements, in response to human behaviors, will help mitigate such occurrences at this interchange.

Introduction

Population growth in the Minneapolis metro area (USA) has been steadily increasing over the past years (U.S. Census Bureau, 2018) and this has increased demand on area infrastructure. Growing traffic volume produces stress on aging highways and interchanges making traffic safety a concern. The Interstate 494 and Interstate 35 West (I-494 and I-35W) cloverleaf interchange, located in the Minneapolis suburb of Bloomington, is an example of this aging interchange under stress.

The cloverleaf structure was designed in the early 20th century to keep traffic flowing without the use of traffic signals (Hotchin, 2017). The cloverleaf design utilizes four loops (ramps) to transfer traffic flow seamlessly from one highway to another while maintaining right-lane exit and entrance flow. However, this creates a juxtapositional situation as traffic interweaves lanes while simultaneously adjusting to decreasing (or increasing) speeds depending on exiting (or entering) pathway. Hotchin (2017) states, "…weaving is a problem that may lead to breakdown in traffic operation and [lead to] more accidents."

The I-494 and I-35W cloverleaf interchange was constructed in the 1960's and has a continued overcapacity daily traffic volume; the daily traffic volume in 2014 was 290,000 vehicles (Minnesota Department of Transportation, 2013). Its safety ranking is among the worst for interchanges in the Minneapolis metro area (Short Elliott Hendrickson, Inc., 2010). Congestion and design flaws on I-494 and I-35W arteries impact the interchange in a shockwave effect (SEH) creating stress on traffic maneuvering through the interchange. During a 10-year span, 2006 through 2015, an average 8.45

Roberts, Candace. 2018. Statistical Analysis of Vehicle Crash Incidents at the Minnesota USA Interstate 494 and Interstate 35 West Interchange using Chi-Square Statistic Goodness-of-Fit Testing. Volume 21, Papers in Resource Analysis. 15 pp. Saint Mary's University of Minnesota University Central Services Press. Winona, MN. Retrieved (date) http://www.gis.smumn.edu

monthly crash incidents were reported as evident in this study's dataset.

The state of Minnesota documentation on traffic collision occurrence preferences the term "crash" to "accident" since accident suggests "a random, unavoidable quality about the events in question" (Minnesota Department of Public Safety, 2016). When, in fact, the reduction in the number and severity of crashes over the past decades is attributed to advances in technology, engineering, public policy, and driver behavior, which points to preventable instances rather than random accounts (MnDPS). Three criteria are required to define a motor vehicle crash: occurrence on a public road, minimum $1000 damage or a person is injured, transport of motor vehicle (MnDPS; Minnesota Department of Public Safety, n.d.).

Understanding impacts leading to vehicle crashes by analyzing historical data provides valuable information in efforts to mitigate such occurrences at the I-494 and I-35W cloverleaf interchange. This study utilizes contingency tables and Chi-Square Goodness-of-Fit tests to determine significant association and statistical equality using weather conditions, driver age, driver behavior, vehicle safety, time, and collision type of 1014 crash incidents from 2006 to 2015 at the I-494 and I-35W cloverleaf interchange.

Methods

Motor vehicle crash occurrences of the I-494 and I-35W interchange from 2006 to 2015 were extensively analyzed using contingency tables to calculate the Chi-Square Goodness-of-Fit tests to determine significant statistical association using single dataset attributes tested for significance with the Chi-Square Goodness-of-Fit test to determine statistical equality.

Data

Historical crash data were obtained from the Minnesota Department of Transportation in shapefile and CSV formats. The data incorporated crash incidents within 1250' from the center of the I-494 and I-35W interchange from 2006 to 2015. The dataset contained 1014 crash incidents involving single and multiple vehicle collisions.

The shapefiles revealed the location for 87% of incidents occurred at the center of the interchange without dispersion of occurrence elsewhere. Either this indicated most likely location for a crash to occur or data recording preference. In either case, location analysis, such as hot spot clustering, were deemed insignificant in determining areas of most likely occurrences due to crash location accuracies. Nonetheless, all 1014 points are shown in Figure 1, of which 881 points are displayed at the same location.

Figure 1. I-494 and I-35W cloverleaf interchange: 1014 motor vehicle crash occurrences, 2006–2015.

Information analyzed for possible

significant statistical association (indicating influence to crash occurrence) were month, day, year, time, collision type, weather conditions, driver contributing factors, and driver age.

Information analyzed for statistical equality testing of occurrence across 2006 to 2015 were crash frequencies, two collision types (rear-end and sideswipe), and two driver contributing factors (driver inattention or distraction, and illegal or unsafe driving speeds).

Data Preparation

Contributing factors due to driver influence, either human behavior or vehicle failure, were recorded in multiple fields as were weather conditions. The fields were equally weighted. Therefore, the fields were analyzed as separate entities, rather than the whole entities as recorded in the crash incident to extract influence of a single entity.

In simplifying weather conditions from the two weather fields, clear was categorized as clear; cloudy or clear/cloudy were categorized as cloudy, reference to rain was categorized as rain, reference to snow was categorized as snow, sleet was categorized as sleet, fog was categorized as fog, and blowing winds were categorized as blowing winds.

Extracting Data Subsets

According to analysis in the Minnesota Motor Vehicle Crash Facts 2015 (Minnesota Department of Public Safety, 2016), driver behaviors are frequently cited as contributing factors to crash incidents more often over vehicular safety failures. This was demonstrably accurate in the *n*=1014 dataset case study for the I-494 and I-35W cloverleaf interchange.

In the *n*=1014 dataset case study, driver behaviors as a contributing factor accounted for 43.5% of the crash incidents, whereas vehicle safety failures were accountable for only 3%. Factors categorized as 'not clearly identified as contributing to crash incident' were 44.3%.

Odds are, if the 44.3% 'not clearly identified factors' were countable as vehicle safety failures, the whole 44.3% would not shift to vehicle safety failures. At most, 22.2% (or half) would shift and the remaining percentage would shift to driver behavior factors, distributing the 44.3% equally; therefore, distribution of factors as 'not clearly identified' were disregarded. With driver behaviors greater than vehicle safety failures, driver behaviors as contributing factors attributed at a higher percentage to crash incidents in the *n*=1014 dataset case study.

The 2015 Minnesota crash analysis (MnDPS) reports, "…most crashes occur in good driving conditions" which included clear weather conditions. The *n*=1014 dataset case study reflected 655 crashes, or 64.6%, occurred during clear weather conditions. The second highest percentage of crash incidents occurred during cloudy weather conditions at 231 crashes (22.8%). Clear or cloudy weather conditions comprised 87.4%. In contrast, precipitous conditions – rain, snow, sleet – accounted for 12% with remaining 0.6% attributed to fog and blowing winds.

A subset from the dataset (based on the results of the majority of crash incidents were due to driver behaviors during clear or cloudy weather conditions) was extracted. The subset contained 402 crash incidents.

The distribution of crashes here from 2006 to 2015 is displayed in scatter plots with the month (as single-letter initials) along the x-axis and the time along the y-axis. Figure 2 is the original dataset case study (*n*=1014); Figure 3 is the data subset (*n*=402). The distributions are similar with saturation of crash incidents during the morning rush hour (6:00 a.m. to 9:00 a.m.), afternoon (12:00 p.m. to 6:00 p.m.), and Thanksgiving through year end; and absence of distribution around April and October. Based on similar distribution, the data subset (*n*=402) was representative of the original dataset case study (*n*=1014).

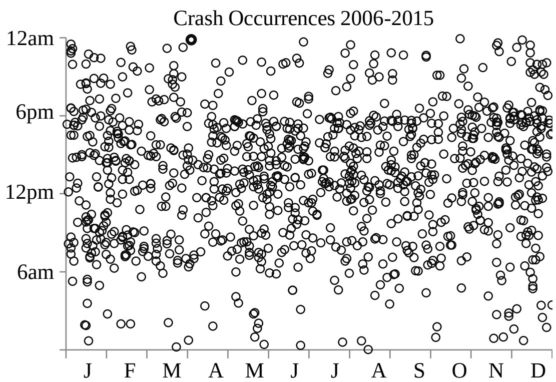

Figure 2. All crash occurrences by month and time from 2006 – 2015 (*n* = 1014). Clustered areas of overlapping, stacking circles are higher densities of crash occurrences.

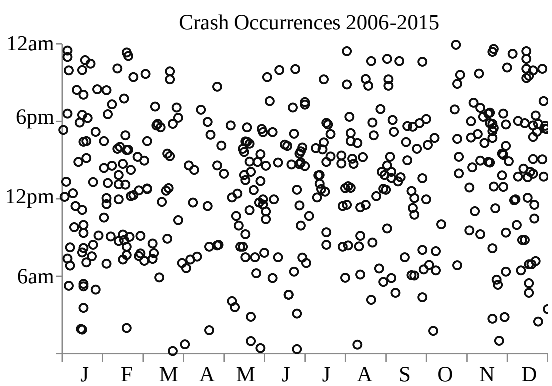

Figure 3. Crash occurrences by month and time from 2006 – 2015 (*n* = 402) due to driver behavior while driving in clear or cloudy conditions.

Contingency Statistical Analysis

The data subset (*n*=402) was analyzed using contingency tables to calculate Chi-Square Goodness-of-Fit to determine significant statistical association as referenced in Zar (2010).

Contingency Tables

A contingency table contains enumeration data of observed frequencies of two categorical variables where the rows represent one variable and the columns represent the other variable. The number of rows is determined by the number of categories within the row variable and the number of columns is determined by the number of categories in the column variable.

In this study (*n*=402), five contingency tables were constructed for analysis. The first contingency table contained two categorical variables, driver behavior and collision types. Driver behaviors were divided into 14 categories (failure to yield, illegal or unsafe driving speed, driving too closely, disregard for traffic control device, improper passing or overtaking, improper or unsafe lane use, improper parking or starting or stopping, improper turn, over-correcting, impeding traffic, driver inattention or distraction, driver inexperience, chemical impairment, and other factor). Collision types were divided into eight categories (rear-end, sideswipe, off-road left side, off-road right side, right angle, head-on, not specified, and other type of collision).

The second contingency table consisted of driver age and driver behavior. Driver age was divided into nine categories (15-19 year-olds, 20-29 year-olds, 30-39 year-olds, 40-49 year-olds, 50-59 year-olds, 60-69 year-olds, 70-79 year-olds, 80-89 year-olds, and not specified). Driver behaviors were divided into 14 categories as in the first contingency table.

The third contingency table incorporated driver age and collision types. Driver age was divided into nine categories as in the second contingency table. Collision types were divided into eight categories as in the first contingency table.

The fourth contingency table comprised of time-periods and driver behavior. Time-periods were divided into eight categories (midnight to 3:00 a.m., 3:00 a.m. to 6:00 a.m., 6:00 a.m. to 9:00 a.m., 9:00 a.m. to noon, noon to 3:00 p.m., 3:00 p.m. to 6:00 p.m., 6:00 p.m. to 9:00 p.m., and 9:00 p.m. to midnight). Driver behaviors were divided into 14 categories as in the first contingency table.

The fifth contingency table contained time-periods and collision types. Time-periods were divided into eight categories as in the third contingency table. Collision types were divided into eight categories as in the first contingency table.

The Chi-Square statistics were derived from the contingency tables and the Goodness-of-Fit tests determined whether to accept or reject the null hypothesis.

Chi-Square Goodness-of-Fit Test

The Chi-Square statistic for evaluating the Goodness-of-Fit was performed on the data subset (*n*=402) to determine significant association between categorical variables. The following Chi-Square statistic formula was performed on data residing in the contingency tables (Zar, 2010):

$$\chi^2 = \sum\sum\frac{(f_{ij} - \hat{f}_{ij})^2}{\hat{f}_{ij}},$$

where f_{ij} is the observed frequency at row i and column j of the contingency table, and $\hat{f}_{ij}$ is the expected frequency at i, j.

The expected frequency was derived from the contingency table using the following formula:

$$\hat{f}_{ij} = \frac{(R_i)(C_j)}{n},$$

where R_i is the total observed frequency of row i, C_j is the total observed frequency of column j, and n is the sample size.

Expected frequency calculations were evaluated using crosstab rules assuring none were less than or equal to 3, and no more than 20% were less than or equal to 5. Columns or rows were combined, if possible, or removed until the expected frequency calculations adhered to the crosstab rules.

Once the crosstab rules were satisfied and the Chi-Square statistic was calculated and assessed using the critical value of the Chi-Square distribution using an alpha (α) level of significance of 0.05 and degrees of freedom calculated as follows:

$$v = (r - 1)(c - 1),$$

where r is the number of rows, and c is the number of columns.

If the Chi-Square statistic was less than the critical value, the null hypothesis was not rejected and that the row and column variables were considered statistically independent ($p > 0.05$).

If the calculated Chi-Square statistic was greater or equal to the critical value, the null hypothesis was rejected ($p < 0.05$) and the contingency table was assessed for possible data reconfiguration – combining rows, columns, or removal – if doing so continued to make the Chi-Square test statistic more statistically significant when the Chi-Square Goodness-of-Fit test was recalculated. If data reconfiguration was not successful, the null hypothesis was rejected ($p < 0.05$)

indicating that a dependent association between variables existed.

Equality Statistical Analysis

The data subset (n=402) was analyzed using the Chi-Square Goodness-of-Fit test to determine significant statistical equality. In this study (n=402), collision type, driver behavior, and time were assessed in eleven Chi-Square tests.

The first test determined the overall significance by evaluating the number of crash occurrences by year.

The second and third tests turned toward evaluating collision types by year. The number of rear-end and sideswipe collisions were examined for statistical significance. Further evaluations continued for significance of rear-end and sideswipe collisions from 6:00 a.m. to midnight generated the fourth and fifth tests.

The sixth and seventh tests veered toward evaluating driver behaviors by year. The number of collisions attributed to driver inattention or distraction, and illegal or unsafe driving speeds were examined for significance. Further evaluations continued for significance of rear-end and sideswipe collisions resulting from driver inattention or distraction, and illegal or unsafe driving speeds created eighth through eleventh tests.

Chi-Square Goodness-of-Fit tests were used to determine whether to accept or reject the null hypothesis.

Chi-Square Statistic Goodness-of-Fit Test

The Chi-Square statistic was performed on the data subset (n=402) to determine statistical equality. The following Chi-Square statistic formula was utilized (Zar, 2010):

$$\chi^2 = \sum \frac{(f - \hat{f})^2}{\hat{f}},$$

where f is the observed frequency, and $\hat{f}$ is the expected frequency.

The expected frequency was derived from observed frequencies using the following formula:

$$\hat{f} = \frac{\sum f}{n},$$

where f is the observed frequency, and n is the sample size.

Once the statistic was calculated, the Goodness-of-Fit test was performed using the critical value of the Chi-Square distribution using an alpha (α) level of significance of 0.05 and degrees of freedom calculated as follows:

$$v = n - 1,$$

where n is the sample size.

If the Chi-Square value was less than the critical value, the null hypothesis was not rejected ($p > 0.05$). This indicated that observed frequencies were statistically equal.

If the calculated Chi-Square statistic was greater or equal to the critical value, the null hypothesis was rejected ($p < 0.05$). Observed frequencies were then subdivided, and Chi-Squares were recalculated to ascertain significant differences between observed and expected frequencies.

Results

Many associations were examined for equality using the Chi-Square Goodness-of-Fit test to determine influences leading to vehicle crashes.

Testing for Association of Factors

Contingency tables were used to calculate the Chi-Square Goodness-of-Fit test to determine significant association in relation to crash occurrences.

Chi-Square Goodness-of-Fit Test

Driver behavior and collision types, driver age and driver behavior, driver age and collision types, time-periods and driver behavior, and time-periods and collision types were assembled into five contingency tables and analyzed for significant statistical association.

Driver Behavior and Collision Types

The data subset (n=402) contained only driver behaviors as contributing factors to crash occurrences, which resided in two fields. Separating and simplifying the fields into single entities for analysis, contributing factors totaled 526 driver behavior involvements in n=402 crash occurrences.

The highest percentages of driver behaviors attributed to crashes were driver inattention or distraction (27.2%), illegal or unsafe driving speed (21.5%), following too closely (20.7%), and improper or unsafe lane use (12%) (Figure 4).

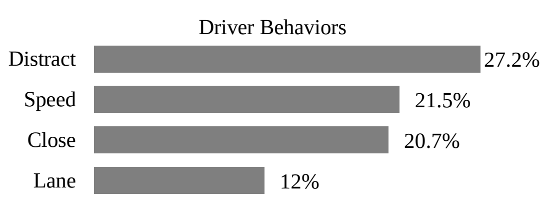

Figure 4. Highest percentages of driver behaviors contributing to crash occurrences (n = 402).

The highest percentages of collision type were rear-end (49.5%), sideswipe (17.2%), off-road left side (12.2%), and off-road right side (8.5%) (Figure 5). Among driver behaviors attributed to those collisions, illegal or unsafe driving speed was the leading cause for off-road left- and right-side collisions; driver inattention or distraction was the second leading cause for rear-end, sideswipe, and off-road left-side collisions.

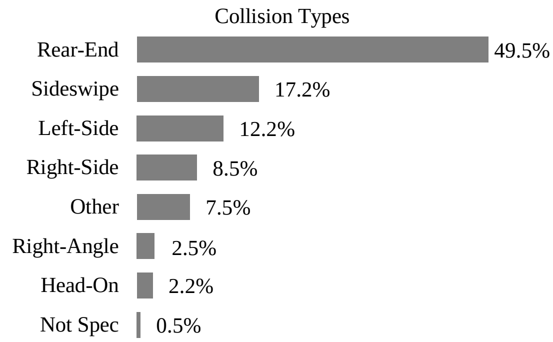

Figure 5. Percentages of collision types contributing to crash occurrences (n = 402).

The Chi-Square Goodness-of-Fit test between driver behavior and collision types led to data reconfiguration into four collision types (rear-end, sideswipe, off-road left side, and off-road right side) and four driver behaviors (driver inattention or distraction, illegal or unsafe driving speed, following too closely, and improper or unsafe lane use). The calculated Chi-Square statistic was significant indicating that a dependent association between driver behavior and collision types existed ($p < 0.05$) as shown in Table 1.

Table 1. Driver behavior and collision types contingency table calculating Chi-Square Goodness-of-Fit test for significant association (n = 402). Expected frequency noted in parentheses.

	Collision Types				
Behavior	Rear-End	Sideswipe	Left-Side	Right-Side	**Total**
Distract	95	24	12	5	136
	(84.34)	*(24.95)*	*(16.17)*	*(10.54)*	
Speed	28	13	27	20	88
	(54.57)	*(16.14)*	*(10.46)*	*(6.82)*	
Close	101	3	1	2	107
	(66.36)	*(19.63)*	*(12.72)*	*(8.29)*	
Lane	16	31	6	3	56
	(34.73)	*(10.27)*	*(6.66)*	*(4.34)*	
Total	240	71	46	30	387

$\chi^2_{0.05,9}$ = 16.919
χ^2 = 170.675; therefore, reject null hypothesis ($p < 0.05$).

Driver Age and Driver Behavior

In the data subset (n=402), the driver age categories were grouped by decade beginning with 20 year-olds; the 15-19 year-olds were grouped as a partial decade.

The highest occurrence of crashes by age group was the 20-29 year-olds at 36.6% which was twice of the next age group (30-39 year-olds) at 18.2% as shown in Figure 6.

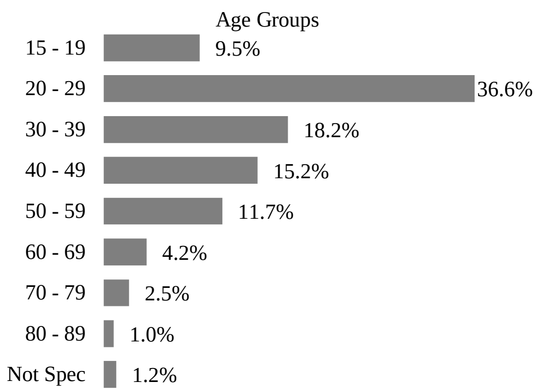

Figure 6. Percentages of age groups contributing to crash occurrences (n = 402).

Although population size within each age group differed, percentage-wise driver inattention or distraction was the leading driver behavior cause in all age groups; and illegal or unsafe driving speed and following too closely were the second or third leading driver behavior cause for age groups 15-59 and third or fourth leading driver behavior cause for age groups 60-89.

The Chi-Square tests for Goodness-of-Fit between age groups and driver behavior led to data reconfiguration by combining age groups 60 and above and four driver behaviors (driver inattention or distraction, illegal or unsafe driving speed, following too closely, and improper or unsafe lane use). The calculated Chi-Square statistic was not significant indicating no association between driver age groups and driver behavior ($p > 0.05$) as shown in Table 2.

Table 2. Driver age and behavior contingency table calculating Chi-Square Goodness-of-Fit test for significant association (n = 402). Expected frequency noted in parentheses.

	Driver Behavior				
Age	Distract	Speed	Close	Lane	**Total**
15-19	15	14	14	1	44
	(14.81)	*(11.47)*	*(11.36)*	*(6.36)*	
20-29	49	49	37	25	160
	(53.84)	*(41.71)*	*(41.33)*	*(23.13)*	
30-39	28	16	20	16	80
	(26.92)	*(20.85)*	*(20.66)*	*(11.56)*	
40-49	21	13	19	7	60
	(20.19)	*(15.64)*	*(15.50)*	*(8.67)*	
50-59	17	12	15	8	52
	(17.50)	*(6.78)*	*(13.43)*	*(7.52)*	
60-89	12	6	4	4	26
	(8.75)	*(6.78)*	*(6.72)*	*(3.76)*	
Total	142	110	109	61	422

$\chi^2_{0.05,15} = 24.996$

$\chi^2 =$ 15.310; therefore, accept null hypothesis ($p > 0.05$).

Driver Age and Collision Types

Across all age groups, rear-end collisions were the leading collision type. Sideswipe collisions were the second leading collision type, except age groups 15-29 (off-road left side) and age group 60-69 (other type of collision). Sideswipe collisions were the third leading collision type for age groups 15-29.

A significant dependent association existed between driver age and collision types. Passing crosstab rules led to consolidating age groups above 50, vehicle-vehicle collision types (rear-end, sideswipe, right-angle, and head-on), vehicle maneuvering off road (off-road left side and off-road right side), and other collision type with collision not specified. The calculated Chi-Square statistic was significant indicating that a dependent association between driver age and collision types existed ($p < 0.05$).

Time-periods and Driver Behavior

In the data subset (*n*=402), the time categories were grouped by three-hour increments. The increment size was chosen to incorporate defined morning and afternoon rush hour time-periods as noted in the Minnesota Motor Vehicle Crash Facts 2015 (Minnesota Department of Public Safety, 2016).

The highest percentage of crashes occurred during the 3:00 p.m. to 6:00 p.m. (afternoon rush hour) with 22.1%. The next highest was noon to 3:00 p.m. at 21.1% as shown in Figure 7. Together, noon to 6:00 p.m., accounted for 43.2% which nearly meets the 2015 statistic of 45% state-wide in this time-period (MnDPS).

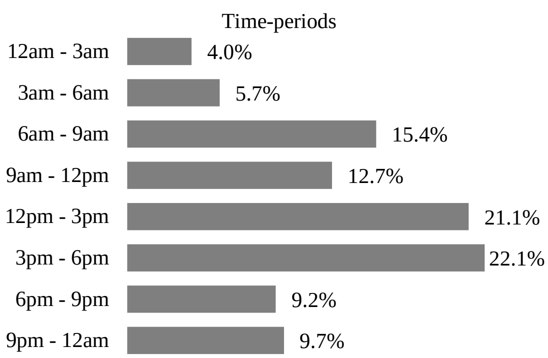

Figure 7. Percentages of time-periods contributing to crash occurrences (n = 402).

Assessing driver behaviors by three-hour increments, driver inattention or distraction was the leading cause for half of the time increments (6:00 a.m. to 9:00 a.m., 9:00 a.m. to noon, noon to 3:00 p.m., and 6:00 p.m. to 9:00 p.m.).

Driver inattention or distraction was the second leading cause from 3:00 p.m. to 6:00 p.m. (the first was following too closely) and 9:00 p.m. to midnight (first was illegal or unsafe driving speed).

Driver inattention or distraction was the third leading cause from midnight to 6:00 a.m. with illegal or unsafe driving speed as the first leading cause and chemical impairment as the second leading cause.

The first Chi-Square Goodness-of-Fit test for significant association between time-periods and driver behavior led to combining time-periods into six-hour increments (midnight to 6:00 a.m., 6:00 a.m. to noon, noon to 6:00 p.m., and 6:00 p.m. to midnight) and reducing driver behaviors to driver inattention or distraction, illegal or unsafe driving speed, following too closely, and improper or unsafe lane use. The calculated Chi-Square statistic here was significant indicating that a dependent association between six-hour increment time-periods and driver behavior existed ($p < 0.05$).

The second Chi-Square Goodness-of-Fit test for significant association between time-periods and driver behavior was a continuation of the first test to determine if it was possible to achieve independence of time-periods and driver behavior. Only when the time-periods were consolidated to 12-hour increments did the calculated Chi-Square statistic achieve no significance indicating no association between time of 12-hour increments and driver behavior ($p > 0.05$), anything less was significant ($p < 0.05$).

Time-periods and Collision Types

The leading collision type from midnight to 3:00 a.m. was off-road left side, 3:00 a.m. to 6:00 a.m. was sideswipe, and 6:00 a.m. to midnight was rear-end collisions. The second leading collision type midnight to 3:00 a.m. was off-road right side, 3:00 a.m. to 6:00 a.m. was off-road left side, 6:00 a.m. to 6:00 p.m. was sideswipe, 6:00 p.m. to 9:00 p.m. was off-road left side, and 9:00 p.m. to midnight was sideswipe collisions.

Time-periods and collision types

Chi-Square Goodness-of-Fit test were shown in two scenario tests. The first test consolidated time-periods to six-hour groups (midnight to 6:00 a.m., 6:00 a.m. to noon, noon to 6:00 p.m., and 6:00 p.m. to midnight), vehicle-vehicle collision types (rear-end, sideswipe, right-angle, and head-on), vehicle maneuvering off road (off-road left side and off-road right side), and other collision type with collision not specified. The Chi-Square statistic calculated here was significant indicating that a dependent association between six-hour increment time-periods and collision types existed ($p < 0.05$).

The second test for time-periods and collision type combined time into 12-hour increments and combined other collision type with collision not specified. Again, the calculated Chi-Square statistic was significant indicating that a dependent association between time of 12-hour increments and collision types existed ($p < 0.05$).

Testing for Equality of Crashes

Collision type, driver behavior, and time were used to calculate Chi-Square Goodness-of-Fit tests to determine significant equality by year.

Chi-Square Goodness-of-Fit Test

Significance of statistical equality was determined by year for overall crash frequencies, rear-end collisions, sideswipe collisions, driver inattention or distraction, and illegal or unsafe driving speeds. These particular collision types and driver behaviors are of interest because of their high frequency of occurrence as shown in Figures 4 and 5.

Additionally, collision types and driver behaviors were plotted on line graphs where yearly totals were connected by lines and trend lines were added to even spikes of extreme variance from year to year.

Collision Types

The two most occurring collision types were rear-end collisions at 49.5% and sideswipe collisions at 17.2% (Figure 5). They were analyzed, alongside the total number of crash occurrences ($n = 402$), for similarities of frequencies and trends (Figure 8).

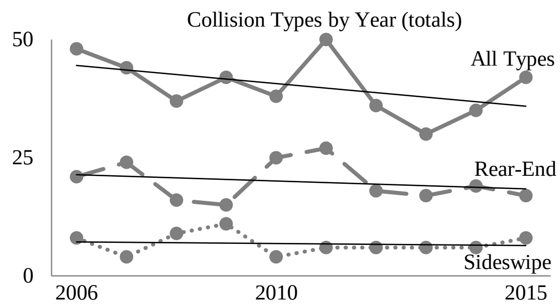

Figure 8. Total crash occurrences by year. All occurrences, rear-end collision types, sideswipe collision types ($n = 402$).

Total number of crash occurrences (n=402) Chi-Square Goodness-of-Fit test revealed statistical equality from 2006 to 2015 ($p > 0.05$), indicating that statistically significant equal number of crashes occurred from year to year as shown in Table 3.

Rear-end collisions (n=402) Chi-Square test revealed statistical equality from 2006 to 2015 ($p > 0.05$), indicating again, a statistically equal number of rear-end collisions occurred from year to year. Since all the rear-end collisions occurred from 6:00 a.m. to midnight, evaluating the Chi-Square Goodness-of-Fit for this specific time-period was identical to the prior results.

Sideswipe collisions (n=402) Chi-Square test revealed statistical equality from 2006 to 2015 ($p > 0.05$). A statistically equal number of sideswipe collisions occurred from year to year. Since 88% of sideswipe collisions occurred from 6:00 a.m. to midnight, whereas 100% of rear-end collisions occurred during this time-period, evaluating Chi-Square Goodness-of-Fit test revealed statistical equality from 2006 to 2015 ($p > 0.05$), indicating a statistically equal number of sideswipe collisions from 6:00 a.m. to midnight occurred from year to year.

Table 3. Yearly crash occurrences Chi-Square Goodness-of-Fit test for significant equality of occurrence (n = 402). Expected frequency noted in parentheses.

	All Crash Occurrences
Year	Frequency
2006	48 *(40.20)*
2007	44 *(40.20)*
2008	37 *(40.20)*
2009	42 *(40.20)*
2010	38 *(40.20)*
2011	50 *(40.20)*
2012	36 *(40.20)*
2013	30 *(40.20)*
2014	35 *(40.20)*
2015	42 *(40.20)*
Total	402

$\chi^2_{0.05,9} = 16.919$

$\chi^2 =$ 8.50; therefore, accept null hypothesis ($p > 0.05$).

Driver Behaviors

The two highest percentages of collisions due to driver behavior were driver inattention or distraction at 27.2% and illegal or unsafe driving speed at 21.5% (Figure 4). They were analyzed, alongside the total number of crash occurrences (n = 402), for similarities of frequencies and trends (Figure 9).

The 27.2% collisions due to driver inattention or distraction, rear-end and sideswipe collisions were analyzed for similarities of frequencies and trends (Figure 10).

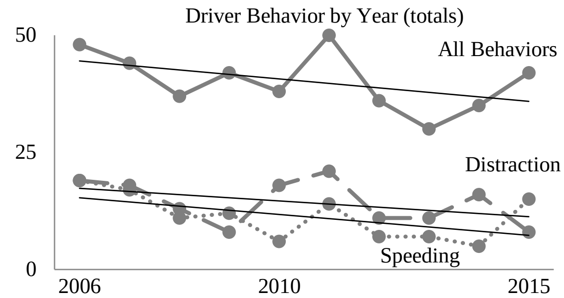

Figure 9. Total crash occurrences by year. All occurrences, distraction driver behavior, speeding driver behavior (n = 402).

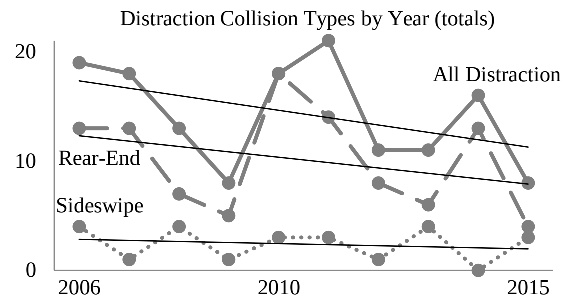

Figure 10. Total crash occurrences by year as influenced by driver distraction. Rear-end and sideswipe collisions due to distraction (n = 402).

Collisions attributed to driver inattention or distraction (n=402) Chi-Square Goodness-of-Fit test revealed statistical equality from 2006 to 2015 ($p >$ 0.05), indicating a statistically equal number of collisions occurred from year to year. Similar results were found for sideswipe collisions attributed to driver inattention or distraction; however, this was not the case for rear-end collisions.

Rear-end collisions attributed to driver inattention or distraction revealed that they were not statistically equal from 2006 to 2015 ($p < 0.05$). Subdividing the observed frequencies and recalculating the Chi-Square statistic goodness-of-fit test revealed that all years were statistically equal ($p > 0.05$), except for year 2010 (Table 4).

Regarding the 21.5% collisions due to illegal or unsafe driving speed, rear-end and sideswipe collisions were analyzed for similarities of frequencies and trends (Figure 11).

Table 4. Yearly rear-end collisions attributed to driver inattention or distraction Chi-Square Goodness-of-Fit test for significant equality of occurrence (n = 402), but with 2010 data removed. Expected frequency noted in parentheses.

Year	Driver Inattention or Distraction Rear-End Collisions
2006	13 *(9.22)*
2007	13 *(9.22)*
2008	7 *(9.22)*
2009	5 *(9.22)*
2011	14 *(9.22)*
2012	8 *(9.22)*
2013	6 *(9.22)*
2014	13 *(9.22)*
2015	4 *(9.22)*
Total	83

$\chi^2_{0.05,8}$ = 15.507
χ^2 = 13.83; therefore, accept null hypothesis ($p > 0.05$).

Collisions attributed to illegal or unsafe driving speeds (n=402) were statistically not equal from 2006 to 2015 ($p < 0.05$). Subdividing the observed frequencies and recalculating the Chi-Square Goodness-of-Fit test revealed that all years were statistically equal ($p >$ 0.05), except for year 2006.

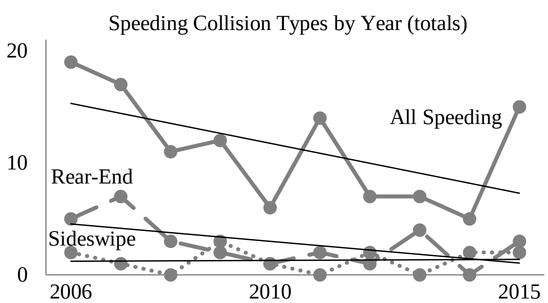

Figure 11. Total crash occurrences by year as influenced by driver speeding. Rear-end and sideswipe collisions due to speeding (n = 402).

Sideswipe collisions attributed to illegal or unsafe driving speeds revealed statistical equality from 2006 to 2015 ($p >$ 0.05), indicating a statistically equal number of collisions occurred from year to year. Similar results were found for rear-end collisions attributed to illegal or unsafe driving speeds.

Discussion

The Chi-Square Goodness-of-Fit test was used to determine association and equality between collision types, driver behaviors, driver age, and time-periods.

Significant dependent association existed with driver behavior and collision types. Rear-end and sideswipe collisions were the top two occurring collision types with commonplace driver behaviors of driver inattention or distraction, and illegal or unsafe driving speed.

Assessing significant equality of occurrence during 2006 – 2015, sideswipe collisions in conjunction with driver inattention or distraction, or illegal or unsafe driving speed were statistically equal. As well, rear-end collisions were statistically equal, except with driver inattention or distraction which occurred statistically higher in 2010.

A significant association existed between driver age and collision types. Investigating further, rear-end collisions were the leading collision type across all driver ages; sideswipe collisions were the second leading collision type for driver ages 30-59 year-olds and 70-89 year-olds – nearly all age groups.

Significant equality of occurrence for rear-end collisions, as well as sideswipe collisions, occurred across all years. Rear-end and sideswipe collisions did not discriminate occurrence by age and occurred in statistically equal numbers from 2006 to 2015.

Significant dependent association existed with time-periods and driver behavior. The leading driver behaviors attributed to 75% of the three-hour time-periods were driver inattention or distraction, and illegal or unsafe driving speed. Regarding significant equality of crash occurrence attributed to driver behavior, all years were statistically equal, except illegal or unsafe driving speed occurred statistically higher in 2006.

A significant dependent association existed between time-periods and collision types. Rear-end and sideswipe collisions accounted for 87.5% of the three-hour time-periods. The time-periods spanning 6:00 a.m. to midnight comprised of the majority of rear-end and sideswipe collisions, and occurred in statistically equal numbers across all years.

The total number of crash occurrences (n = 402) by year showed a downward sloping trend line indicating crash occurrence frequency decreasing (Figure 8). Comparing crash occurrences attributed to rear-end collisions to all types of collisions, the trend line slopes are somewhat similar as does the frequencies of occurrence by year (Figure 8). The similarities indicate possible relationship between rear-end collisions and all collisions, and furthermore extrapolate that most collisions are rear-end collisions as shown in Figure 5 of percentages of collision types contributing to crash occurrences.

Again, comparing crash occurrences attributed to sideswipe collisions to all types of collisions, the trend line slopes are somewhat similar, although the sideswipe trend line has less slope (flatter) as shown in Figure 8. The similarities indicate a possible relationship, but not as strong as the relationship between rear-end collisions and all types of collisions. Sideswipe collisions were the second leading collision type as shown in Figure 5; the flatter trend line is evident in the 32.3% difference between rear-end and sideswipe collision occurrence (Figure 5).

Comparing the total number of crash occurrences (n = 402) by year with the crash occurrence frequencies of driver behaviors (Figure 9) is remarkably similar to collision types (Figure 8). Trend lines for collisions attributed to driver inattention or distraction, and illegal or unsafe driving speed have similar downward slope rates as well frequencies by year (Figure 9). Trend line slope similarities are reflected by the 5.7% difference as first and second highest contributing factors to occurrences (Figure 4).

The total number of crash occurrences resulting from driver inattention or distraction (n = 402) by year showed a downward sloping trend line indicating crash occurrence frequency decreasing (Figure 10). Comparing with the occurrences resulting in rear-end collisions from driver inattention or distraction, the trend line slopes are somewhat similar (Figure 10), indicating possible relationship between driver inattention or distraction and those resulting in rear-end collisions. However, dissimilar extrapolations made from comparing all occurrences from driver inattention or distraction with those resulting in sideswipe collisions. Here, the trend line slope of sideswipe collisions has less slope (flatter) indicating less relationship between driver inattention or distraction and those resulting in sideswipe collisions.

The total number of crash occurrences resulting from illegal or unsafe driving speed (n = 402) by year showed a downward sloping trend line indicating crash occurrence frequency decreasing (Figure 11). Comparing with the occurrences resulting in rear-end collisions from illegal or unsafe driving speed, the trend line slopes are somewhat similar (Figure 11), indicating possible relationship between illegal or unsafe driving speed and those resulting in rear-end collisions. The trend line from occurrences resulting in sideswipe collisions is flat (nearly a slope rate of zero) which is dissimilar to the trend line of all occurrence from illegal or unsafe driving speed indicating less relationship between from illegal or unsafe driving speed and those resulting in sideswipe collisions.

Statistically higher occurrences of crash incidents due to driver behavior in 2006 (illegal or unsafe driving speed) and 2010 (rear-end collisions caused by driver inattention or distraction) could have been influenced possibly by better weather driving conditions, for which drivers tend to drive faster (Minnesota Department of Public Safety, 2016).

Events in the years between 2006 and 2010 may have decreased influence in 2007 – 2009, resulting in 2006 and 2010 appearing as anomalies of higher influence. The economic downturn from the Great Recession (2007 – 2009) may have affected driving behavior as unemployment grew and recovery outlook dimmed (Streff, 2001). Highway construction in 2008 – 2009 along arteries surrounding the I-494 and I-35W interchange may also have affected traffic flow in the area.

Further research into influences such as weather, economic, construction, convention and social events, and personal technology devices from a historical perspective would be beneficial in understanding the underlying influences impacting driver behavior leading to crash incidents from 2006 to 2015 at the I-494 and I-35W interchange.

Conclusion

The I-494 and I-35W cloverleaf interchange safety ranking is among the worst for interchanges in the Minneapolis metro area (Short Elliott Hendrickson, Inc., 2010). As analysis in this study validates, the majority of crashes from 2006 through 2015 occurred during good weather conditions which suggests modification to driver behavior is highly desirable, especially considering driving challenges of inclement weather conditions were shown as noninfluential factors to crash occurrence.

Targeting driver behavior modification efforts to reduce motor vehicle crashes in the I-494 and I-35W cloverleaf interchange is highly recommended. Effective educational objectives addressing the driver behaviors of inattention, distraction, and speeding are suggested for drivers of all ages, driving during any hour of the day.

Data Limitations

A remarkable 87% of the data points are located in the center of the interchange. This seemingly indicates the most likely location for crash occurrence or, perhaps, a data recording preference.

Cross-referencing data point locations with other descriptive attributes about crash occurrences (such as roadway direction, character, design; vehicle travel direction; and collision type) did not indicate an assurance to center interchange location. Nor did the descriptive attributes provide a general sense of occurrence placement to confidently adjust location.

Positional accuracy would have opened opportunities to spatial hot spot and regression analysis for areas of most likely occurrence within the interchange. Positional accuracy also would have led to 2- and 3-dimensional mapping exploration for visual analysis and comprehension.

Statistical Analysis Concerns

Results and conclusions were based on the data presented. Possible chance of error may occur in analysis at any point which materializes a false conclusion. Additional information enhancing the data may influence analysis and conclusions likewise. As always, additional studies over more years of data along with a study of variables and conditions impacting driving would be highly desirable.

Acknowledgements

I thank Mr. Derek Leuer and Ms Kitty Hurley for access and speedy delivery of site I-494 and I-35W crash data for which made this study possible.

The honor of Dr. David McConville as advisor and professor – always available, realistic, and kindly patient – and Mr. John Ebert as co-advisor and professor – always quick with enthusiasm and encouragement – were appreciated.

Ms Greta Poser, professor, whose patient expertise assures solid GIS foundation. Saint Mary's University of Minnesota GIS Department where diverse learning environment built a rich learning community was always noted and appreciated, too.

Thanks to my husband and son for allowing me to cocoon into studies for years – life will return to normal*...soon!* And, to my parents for instilling the importance of education – thank you.

References

Hotchkin, S. 2017. The amazing world of: Interchange designs (sightseeing series). Retrieved from http://www.sehinc.com/news/amazing-world-interchange-designs.

Minnesota Department of Public Safety. n.d. In case of a crash: A guide for Minnesota motorists. Retrieved from https://dps.mn.gov/divisions/ots/educational-materials/Documents/In-Case-of-Crash-Brochure.pdf.

Minnesota Department of Public Safety. 2016. Minnesota motor vehicle crash facts 2015. St. Paul, MN: Office of Traffic Safety.

Minnesota Department of Transportation. 2013. I-494/I-35W interchange: Fact sheet 1 – interchange history & need. Retrieved from http://www.dot.state.mn.us/metro/projects/i494and35winterchange/pdf/interchangehistoryneed.pdf.

Short Elliott Hendrickson, Inc. 2010. Preliminary design report I-494/I-35W interchange preliminary (seh no. mntco 107371). Minnetonka, MN: Author.

Streff, E. (Ed.). 2001, June. Connecting the Minnesota safety agenda: Towards zero deaths. Minneapolis, MN: Center for Transportation Studies, University of MN.

U.S. Census Bureau, Population Division. 2018. Annual estimates of the resident population: April 1, 2010 to July 1, 2017. Retrieved from https://factfinder.census.gov/faces/tableservices/jsf/pages/productview.xhtml?src=bkmk.

Zar, J. 2010. *Biostatistical analysis*, 5th ed., pp. 466-516. Upper Saddle River, NJ: Prentice Hall, Inc.