Journal of Safety Research, Vol. 18, pp. 121-136, 1987

0022-4375/87 $3.00 + .00
Printed in the USA

An Analysis of Truck-Involved Freeway Accidents Using Log-Linear Modeling

Thomas F. Golob and Wilfred W. Recker

Log-linear modeling is advanced as a procedure to identify factors that underlie the relative frequency of occurrence of various accident characteristics, such as accident type, location, and severity. The method is centered on the estimation of saturated log-linear models for pairs of accident variables and determination of indices of association between categories of the variables. Using data drawn from more than 9,000 truck-involved accidents that occurred over a 2-year period on freeways in three metropolitan counties in Southern California, the method is demonstrated by analyzing accident characteristics both by type and by freeway route segment. Accidents by collision type are analyzed relative to characteristics such as the primary collision factor, the location of the accident, the time period, road conditions, and weather conditions. Differences among 38 specific freeway segments in terms of accident characteristics are also analyzed. The results of the analyses indicate that the method is a useful tool in uncovering underlying patterns in accident characteristics.

The research reported here involves statistical analyses of the characteristics of truck-related accidents that occurred on freeways in three metropolitan counties in Southern California. It is part of a larger study of the congestion costs associated with truck-related freeway accidents (Golob, Recker, & Leonard, 1987). The analyses are based on data for more than 9,000 truck-involved accidents that occurred during the 1983–1984 period. These data were drawn from the Traffic Accident Surveillance and Analysis System (TASAS) data base maintained by the California Department of Transportation (Caltrans).

The methodology used in this research is centered on the discrete multivariate method of log-linear modeling. The analysis involved the estimation of saturated log-linear models for pairs of accident variables, followed by calculation of indices of association between categories of the variables.

Thomas F. Golob is a research associate at the Institute of Transportation Studies, and Wilfred W. Recker is director of the Institute of Transportation Studies and professor at the Department of Civil Engineering, University of California, Irvine, California 92717.

The reported research was sponsored in part by the AAA Foundation for Traffic Safety and the California Department of Transportation. The authors wish to thank Messrs. Donald Juge and Goro Endo of Caltrans District Seven, Los Angeles, for their support and encouragement. Keith Rutherford prepared the graphics, and Cynthia Wenks prepared the manuscript. The authors alone are responsible for any errors of omission or commission.

TABLE 1
VARIABLES USED IN ANALYSIS

Variable	Categories	Frequency (N = 9,508)
Collision Type	Sideswipe	4,092
	Rear-end	2,964
	Broadside	456
	Hit object	1,108
	Overturn	272
	All other types	616
Primary Collision Factor	Influence alcohol	353
	Tailgating	263
	Failure to yield	65
	Improper turn	903
	Speeding	2,786
	Other violations (hazardous)	4,276
	Other improper driving	189
	Not driver	525
	Unknown	136
Generic Location	Highway	7,889
	Ramp (includes connectors)	1,619
Time Period	midnight - 5:59 a.m.	669
	6:00 a.m. - 8:59 a.m.	1,613
	9:00 a.m. - 11:59 a.m.	2,039
	noon - 2:59 p.m.	2,127
	3:00 p.m. - 5:59 p.m.	1,871
	6:00 p.m. - 8:59 p.m.	728
	9:00 p.m. - 11:59 p.m.	438
Terrain	Flat	8,057
	Rolling	904
	Mountainous	547
Road Conditions	No unusual conditions	9,030
	Holes or loose material	76
	Construction	253
	Other unusual conditions	111
Weather	Clear	7,415
	Cloudy	1,327
	Rain or fog	749
Road Surface Condition	Dry	8,423
	Wet	987
	Icy or otherwise slippery	63
Ramp Direction[a]	On-ramp	581
	Off-ramp	991
	Other (scales, etc.)	47
Ramp Location[a]	Ramp intersection (exit)	451
	Ramp	520
	Ramp entry	220
	Intersecting street	419

[a] Ramp accidents only.

The analyses are divided into two categories: (a) accident characteristics by type of collision and (b) accident characteristics by freeway route segment. In each category, the objective is to identify underlying patterns in accident characteristics. First, accidents by collision type are analyzed relative to characteristics such as the primary collision factor, the location of the accident, the time period, road conditions, and weather. Next, statistical models are developed that identify differences among freeway segments in terms of accident characteristics. Thirty-eight specific freeway segments in Southern California are analyzed as an example of how the method can be used to identify roadways with varying accident characteristics.

DATA

The data source was the TASAS data base maintained by the California Department of Transportation (Caltrans, 1978). This data base comprises all accidents on the state highway system that involved police investigations at the scene of the accident. For 1983–84, there were 9,508 such accidents involving trucks larger than pickups or panel trucks on 22 freeway routes in Los Angeles, Orange, and Ventura Counties, three adjacent metropolitan counties in Southern California.

The analysis focused on the variables listed in Table 1. All variables are categorical in that there is no preconceived ordering of the categories. The category frequencies are included in Table 1. The overall sample size of 9,508 truck-involved accidents over 2 years was sufficient to satisfy minimum cell size requirements in the cross-classifications of most variable pairs. A general rule for the accuracy of the statistical measures used is that all cells (category pairs) in a cross-classification have at least one observation, and 80% of the cells have at least five observations (Cochran, 1954; Haberman, 1978, Vol. I). This was satisfied except in a few situations, which are indicated in the descriptions of the results.

Freeway design, traffic levels, and many other factors that can be broadly defined as freeway conditions are expected to influence the characteristics of truck-involved accidents. The approach in the present study was to divide freeway routes in the case study region into segments, with the conditions within each segment being relatively homogeneous compared to differences in conditions between the segments. The number of possible segments is limited by the necessity of having sufficient numbers of accidents to conduct reliable statistical analyses. Of the 22 freeway routes, 16 had sufficient numbers of accidents. With the aid of the California Department of Transportation (Caltrans), 38 freeway segments were defined on these 16 routes. These 38 segments are mapped in Figure 1, and descriptions of the segments are provided in Table 2.

METHOD

Log-linear models were used to determine relationships between the categorical variables measuring the characteristics and locations of truck-involved freeway accidents. The variables analyzed included the type of collision, the seven other accident characteristic variables listed in Table 1, and the route segment.

Log-linear models are designed to identify structure in the relationship between two or more categorical variables. In the following, the relationship between freeway route segment and collision type is used as an example in describing the modeling approach. The objective in this example is to determine whether or not there are differences among the types of collision that occur on specific route segments. Given that a certain number of truck-involved accidents occur on a specific segment, and that there is a known distribution among types of collisions for all segments, is there a significant interaction between route segment and collision type that indicates that the distribution of collision types might be different for the segment in question? The approach to this question involves estimating a saturated log-linear model for the contingency table represented by the cross-tabulation of route segment by collision type, a 38 × 6 contingency table.

A test of independence between route and accident type involves whether or not the entries in the contingency table can be considered the result of a random process that depends only on the expected number of accidents for each route (for all types) and the expected number of accidents by collision type (for all routes). Defining n_{ij} = observed number of accidents of type j on route i, the hypothesis of independence between route and type involves comparison of each n_{ij} with the randomly expected numbers, m_{ij}, given by the product of the sample size times the probability that an observation falls into the i^{th} row times the probability that the observation falls into the j^{th} column:

$$1.\quad m_{ij} = N(n_{i.}/N)\,(n_{.j}/N) = n_{i.}n_{.j}/N$$

where

$n_{i.}$ = total accidents of all types on route i,
$n_{.j}$ = total accidents of type j on all routes, and
N = total accidents (size of the sample)

The most common measure of association between n_{ij} and m_{ij} is given by:

$$2.\quad \chi^2 = \sum_{ij} (n_{ij} - m_{ij})^2/m_{ij}$$

which has the known chi-square distribution for hypothesis testing under the usual assumption of multinomial distributions and sufficient expected cell frequencies.

Taking the natural logarithm of both sides of Equation 1,

$$3.\quad \ln m_{ij} = \ln N + \ln(n_{i.}/N) + \ln(n_{.j}/N)$$

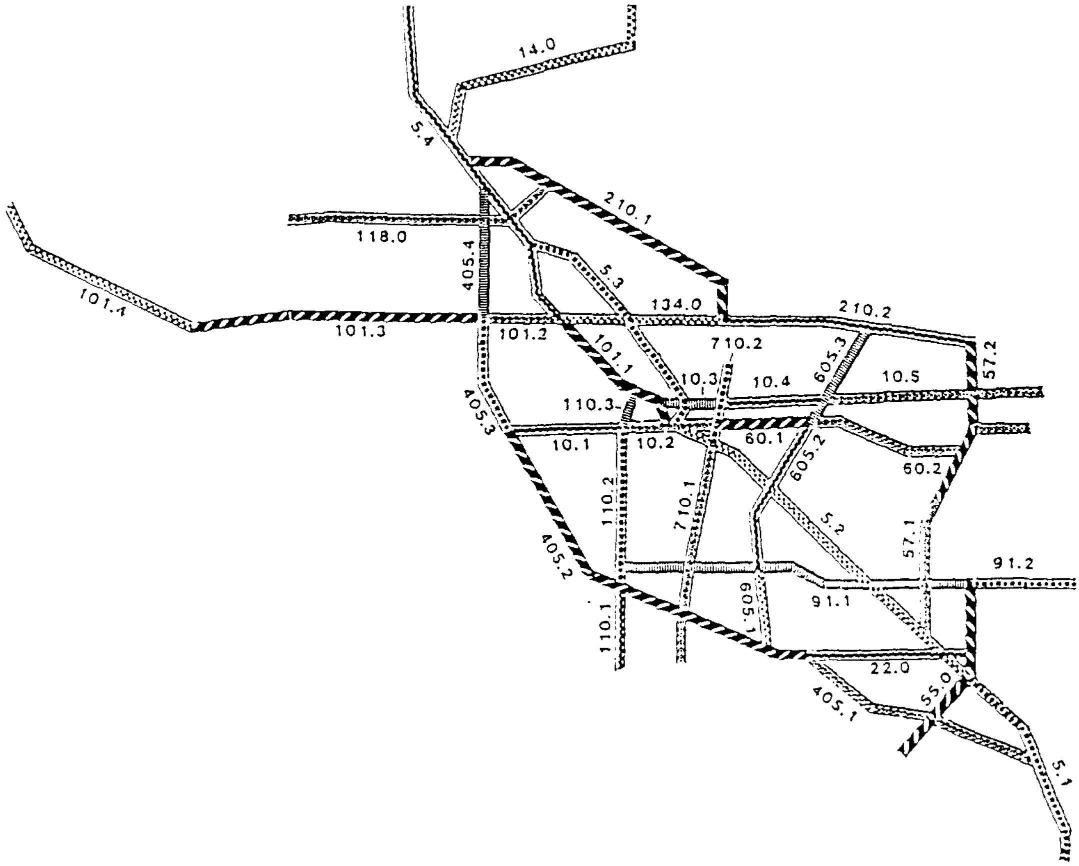
the test of independence for the (i,j) cell of the contingency table translates into a test of whether or not there is a statistically significant β_{ij} term in the log-linear equation

$$4.\quad \ln n_{ij} = \alpha + \beta_i + \beta_j + \beta_{ij}$$

where α accounts for the sample size (grand mean), β_i accounts for the route effect, β_j accounts for the accident type effect, and β_{ij} represents the interaction between route i and type j.

It is logical that the probability process

FIGURE 1
THIRTY-EIGHT FREEWAY SEGMENTS



underlying the accident counts is Poisson. The usual assumption for stochastic processes of Equation 4 is then assumed to include a Poisson error term and represents a saturated log-linear model (Birch, 1963; Plackett, 1962). (Extensive overviews of general families of such models are provided in Bishop, Fienberg, & Holland, 1975; Goodman, 1972, 1978; Haberman, 1974, 1978; McCullagh & Nelder, 1984; Plackett, 1974).

Estimation of the parameters of Equation 4 and their error terms is effectively accomplished using maximum likelihood methods (Bock & Yates, 1973; Haberman, 1973a; Nelder & Wedderburn, 1972). *T*-statistics, given by the ratios of the β_{ij} parameter estimates to the standard errors of the estimates, are used to determine which of the combinations of route (i) and accident type (j) have interaction terms that are significantly different from zero under the assumption of Poisson distributions.

The log-likelihood ratio statistic, given by

$$5. \quad L^2 = 2 \sum_{ij} n_{ij} \log(n_{ij}/m_{ij})$$

has a distribution that is asymptotically chi-square (Cochran, 1954; Haberman, 1978) and can be used to test the hypothesis that the structure of the contingency table can be represented by a log-linear model with some coefficients set to zero.

A direct measure of the degree to which any route-accident type combination (in general, any cell i,j in a contingency table) varies from its expected value is given by the standardized chi-square residual for the cell:

$$6. \quad r_{ij} = [n_{ij} - (n_{i.}n_{.j}/N)] \,/\, \sigma_{ij}$$

where σ_{ij} is the standard error for the cell, given by

$$7. \quad \sigma_{ij} = (n_{i.}n_{.j}/N)^{\frac{1}{2}}$$

This residual is distributed as a standard normal variate under the probability assumptions and sufficient cell frequencies (Haberman, 1973b). The residuals are employed in the present analyses as indices of variation from expected values. They are listed for variable combinations (or interaction terms) that have significant coefficients in the log-linear models. They are not residuals associated with the fits of the log-linear models, which are exact because there are as many parameters as there are cells in the contingency tables ("saturated" models). The standardized residuals merely are one measure of the degree of variance between

TABLE 2
DESCRIPTIONS OF 38 FREEWAY SEGMENTS

Code	Description
5.1	Santa Ana (I-5): Orange-San Diego County line to Jct. 55 (Costa Mesa Freeway)
5.2	Santa Ana (I-5): Jct. 55 to Jct. 10/60 (Pomona Freeway)
5.3	Santa Ana-Golden State (I-5): Jct. 10/60 to Jct. 170 (Hollywood Freeway)
5.4	Golden State-Hollywood (SR-170, I-5): Jct. 101/134 to Jct. 170/5 to Los Angeles/Kern County line
10.1	Santa Monica (I-10): Jct. 405 (San Diego Freeway) to Jct. 110 (Harbor Freeway)
10.2	Santa Monica (I-10)-Pomona (SR 60): Jct. 110 to Jct. 710 (Long Beach Freeway)
10.3	San Bernardino (I-10): Jct. 101 to Jct. 710 (Long Beach Freeway)
10.4	San Bernardino (I-10): Jct. 710 to Jct. 605 (San Gabriel R. Freeway)
10.5	San Bernardino (I-10): Jct. 605 to Los Angeles-San Bernardino County line
14.0	Antelope Valley (SR 14): Begin Jct. 5 (Golden State Freeway) to Los Angeles-Kern County line
22.0	Garden Grove (SR 22): Jct. 405 (San Diego Freeway to end, Jct. 55 (Costa Mesa Freeway)
55.0	Costa Mesa (SR 55): Begin Freeway southwest of 73 to end, Jct. 91 (Riverside Freeway)
57.1	Orange (SR 57): Begin Jct. 5/22 to Orange-Los Angeles County line
57.2	Orange (SR 57)-Pomona (SR 60)-Foothill (I-210): County Line to Jct. 30
60.1	Pomona (SR 60): Jct. 710 (Long Beach Freeway) to Jct. 605 (San Gabriel R. Freeway)
60.2	Pomona (SR 60): Jct. 605 to Los Angeles-San Bernardino County line (excluding overlap with Rte. 60)
91.1	Artesia-Redondo Beach-Riverside (SR 91): Begin Freeway near Jct. 110 (Harbor Freeway to Jct. 55)
91.2	Riverside (SR 91): Jct. 55 to Orange-San Bernardino County line
101.1	Santa Ana-Hollywood (US 101): Begin Jct. 5 (Golden State Freeway) to Jct. 134/170
101.2	Ventura (US 101): Jct. 134/170 to Jct. 405 (San Diego Freeway)
101.3	Ventura (US 101): Jct. 405 to Los Angeles-Ventura County line
101.4	Ventura (US 101): Los Angeles-Ventura County line to Ventura-Santa Barbara County line
110.1	Harbor (I-110): Begin Freeway near Jct. 47 to Jct. 405 (San Diego Freeway)
110.2	Harbor (I-110): Jct. 405 to Jct. 10 (Santa Monica Freeway)
110.3	Harbor (I-110): Jct. 10 to Jct. 101 (Hollywood Freeway)
118.0	Simi Valley-San Fernando Valley (SR 118): Begin Freeway in Ventura County to Jct. Rte. 210
134.0	Ventura (SR 134): Jct. 101/170 (Hollywood Freeway) to Jct. 210 (Foothill Freeway)
210.1	Foothill (I-210): Begin Jct. 5 (Golden State Freeway) to Jct. 134 (Ventura Freeway)
210.2	Foothill (I-210): Jct. 134 to end, Jct. 30
405.1	San Diego (I-405): Begin Jct. 5 (Santa Ana Freeway) to Jct. 22 (Garden Grove Freeway)
405.2	San Diego (I-405): Jct. 22 to Jct. 10 (Santa Monica Freeway)
405.3	San Diego (I-405): Jct. 10 to Jct. 101 (Ventura Freeway)
405.4	San Diego (I-405): Jct. 101 to end, Jct. 5 (Golden State Freeway)
605.1	San Gabriel River (I-605): Begin Jct. 22 to Jct. 91 (Artesia Freeway)
605.2	San Gabriel River (I-605): Jct. 91 to Jct. 60 (Pomona Freeway)
605.3	San Gabriel River (I-605): Jct. 60 to end, Jct. 40 (Foothill Freeway)
710.1	Long Beach (I-710): Begin Jct. 1 to Jct. 5 (Santa Ana Freeway)
710.2	Long Beach (I-710): Jct. 5 to break in route, Valley Blvd., north of 10

actual counts and counts expected under the assumption of independence between the variables.

The log-linear models for this example, as well as for the remaining associations tested in this analysis, were implemented using the GLIM (Generalized Linear Interactive Modeling) program (Baker & Nelder, 1978; McCullagh & Nelder, 1983; Nelder & Wedderburn, 1972). (Log-linear models are also available in most commonly used statistical analysis packages such as SAS, SPSS-X, and BMDP.) The results of these analyses are described below.

RESULTS

Accident Characteristics by Collision Type

Primary collision factor. The relationship between collision type and primary collision factor was analyzed by estimating a log-linear model on the 54-cell cross-classification table for these two variables. The chi-square statistic (Equations 1 and 2) for this cross-classification ($\chi^2 = 4925.4$, $df = 40$) indicated a very strong relationship between the variables (the critical chi-square value being 55.8 at the $p = .05$ level). The log-linear model (Equation 4) for the table had 35 individual cell terms (β_{ij}) that were significantly different from zero at the $p = .05$ level. The standardized residuals (Equations 6 and 7) are listed for these 35 cells in Table 3. (In Table 3 and all subsequent tables, standardized residuals are shown only for cells with significant log-linear model coefficients; all other cells are left blank.)

The residual values in Table 3 indicate relationships that are largely as expected. However, they do reveal some associations that can be useful in explaining accident causality. For instance, rear-end collisions had a strong relationship not only with tailgating driving behavior, but also with alcohol, speeding, and other improper driving. The strongest associations were for speeding (positively associated with rear-end collisions, negatively associated with sideswipes), other violations (positively associated with sideswipes, negatively associated with rear-end collisions), and the not-driver factor (positively associated with other types of collisions).

Accident location. There were substantial differences among collision types in terms of proportional occurrences at highway and ramp locations ($\chi^2 = 508.7$, $df = 5$, critical value = 11.1), again indicating strong interactions. The residuals shown in Table 4 for significant terms in the log-linear model reveal that all collision types except those in the "other" category have had varying highway versus ramp splits, with rear-end and sideswipe collisions occurring predominately at highway sites, and overturns, broadsides, and hit-objects occurring at ramp sites. The strongest associations between collision type and site were for overturns and broadside collisions at ramp locations.

Time period. Collision type and time period (in seven categories) were strongly related ($\chi^2 = 186.1$, $df = 30$, critical value = 43.8). Significant time-of-day patterns for the collision types are shown in Table 5. Hit-object collisions tended to occur from midnight to 6:00 a.m., whereas sideswipes did not. Rear-end collisions appeared to be particularly a morning rush hour phenomenon, and overturns occurred more frequently than expected during the 9:00 p.m. to midnight period. The strongest association involved the occurrence of hit-object collisions during the midnight to 6:00 a.m. period. There were no significant differences among the collision types in terms of their occurrences over days of the week.

Roadway terrain. The relationships between collision types and roadway terrain are shown in Table 6. There was a highly significant overall relationship between the variables ($\chi^2 = 101.5$, $df = 10$, critical value = 18.3). Sections of Routes 14, 405, 118, and 5 are classified in the TASAS highway records as being "mountainous," and almost all routes have both "flat" and "rolling" sections. As shown in Table 6, only the mountainous terrain exhibited differences in the distribution of collision types. There were relatively more rear-end and overturn collisions and relatively fewer sideswipes on mountainous sections.

TABLE 3
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY PRIMARY COLLISION FACTOR

Primary Collision Factor	Collision Type					
	Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
Influence alcohol	-4.2	+6.2				-2.9
Tailgating	-10.0	+16.9	-2.4	-3.0	-2.4	-3.6
Failure to yield		-4.3	+19.2	-2.8		
Improper turn	+9.2	-11.1	+2.7	+3.8	-2.3	-5.4
Speeding	-25.9	+29.8		+3.2	+7.1	-5.7
Other violations	+24.2	-21.5		-8.5	-5.9	
Other improper driving	-2.8	+2.8				
Not driver	-10.1	-8.6		+13.9	+7.7	+22.2
Unknown		-2.5				

Road conditions. There was also a significant relationship between collision type and road conditions ($\chi^2 = 75.9$, $df = 15$, critical value = 25.0). As shown in Table 7, hit-object collisions were more prevalent in areas of construction or other unusual conditions. Collisions in the "other" category occurred in areas classified as having holes or loose material, and this is the strongest association in the table.

Weather conditions. With regard to weather conditions (Table 8), broadside, hit-object, and other types of collisions occurred relatively more frequently in conditions of

TABLE 4
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY LOCATION: HIGHWAY VERSUS RAMP

Accident Location	Collision Type					
	Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
Highway						
Ramp	-4.5	-7.3	+10.5	+8.6	+12.7	

TABLE 5
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH(+) OR LOW (-) FREQUENCIES BY TIME PERIOD

Time Period	Collision Type					
	Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
midnight - 5:59 a.m.	-5.5			+7.3		
6:00 a.m. - 8:59 a.m.		+3.5		-3.8		
9:00 a.m. - 11:59 a.m.						
noon - 2:59 p.m.						
3:00 p.m. - 5:59 p.m.			-2.2	-3.5		
6:00 p.m. - 8:59 p.m.						
9:00 p.m. - 11:59 p.m.					-2.7	

rain or fog, and these are the strongest associations in the table. Conversely, sideswipe collisions were less likely to occur during rainy or foggy conditions. The overall relationship between collision type and weather was again highly significant ($\chi^2 = 201.3$, $df = 10$, critical value = 18.3).

Surface conditions. Table 9 shows a significant relationship between collision type and surface condition ($\chi^2 = 248.0$, $df = 10$, critical value = 18.3). Hit-object and other collisions occurred relatively more often under both wet and icy or slippery road surface conditions. However, broadsides were related to wet roads only, and overturns were related to icy or slippery conditions. The largest standardized deviations from randomly expected frequencies were associated with the occurrences of truck-involved hit-object and broadside collisions on wet freeways.

TABLE 6
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH(+) OR LOW (-) FREQUENCIES BY TERRAIN AT SITE

Terrain	Collision Type					
	Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
Flat						
Rolling						
Mountainous	-6.5	+4.8			+4.1	

TABLE 7
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY ROAD CONDITIONS

Road Conditions	Collision Type: Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
No unusual conditions						
Holes or loose material	-2.2					+4.1
Construction	-3.1			+3.2		
Other unusual conditions				+3.6		

TABLE 8
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY WEATHER CONDITIONS

Weather	Collision Type: Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
Clear				-3.1		
Cloudy		-2.8				
Rain or fog	-6.2		+7.2	+7.6		+3.8

TABLE 9
STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY ROAD SURFACE CONDITIONS

Road Surface Conditions	Collision Type: Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
Dry				-3.2		
Wet	-6.4		+7.7	+8.2		+2.7
Icy or otherwise slippery	-3.1			+4.3	+3.1	+2.4

Accident Characteristics by Freeway Segment

Log-linear models can be used to identify roadway segments with varying accident characteristics. This study presents an example of the types of information that may be provided by this approach. Using 38 freeway segments in Southern California (Figure 1), the analyses focus on several accident characteristics: collision type, relative concentration of ramp involvement, entry vs. exit incidents, location on the ramp, and time of occurrence. These analyses seek to identify freeway segments that tend to have either a particularly severe or a light association with the various accident characteristics.

The relationship between freeway segment (38 categories) and collision type (6 categories) was significant ($\chi^2 = 558.8$, $df = 185$, critical value $= 224.6$). (Although there are 228 cells in the cross-classification of these two variables, only 37 of these cells or 16.2% had expected frequencies of fewer than five accidents, so the chi-square statistic is a fairly accurate indication of overall association.) There were 34 significant interaction terms in the log-linear model of the cross-classification, and the standardized residuals for the category combinations corresponding to these terms are given in Table 10. Because collision type captures an array of other accident characteristics, as described in the previous section, the results in Table 10 are depicted in the tables by collision type.

Sideswipe collisions. The freeway segments with proportionally higher concentrations of sideswipe collisions were segments 10.2 and 5.3 (Table 11). These two adjacent segments are highly congested and serve downtown Los Angeles. The segments with relatively low concentrations of sideswipes were 5.4, 14.0, 605.7, and 57.1. The first two of these segments are located at the northern edge of the metropolitan area, and congestion levels on all four of these segments were substantially lower than the average for all segments.

The indicated relationship between the proportion of sideswipe accidents and traffic congestion was further investigated through correlation analyses of annual average daily traffic measures at locations along each of the 38 freeway segments. The best indicator of the percentage of sideswipe accidents was found to be the maximum annual average daily traffic (AADT) per lane at all locations along a freeway segment; the correlation between this indicator and the percentage of sideswipe collisions was 0.44 ($p < .01$). The maximum AADT per lane for the two segments with high sideswipe incidences was 105,500 for the segment 10.2 and 91,500 for the segment 5.3. The maximum AADT per lane for each of the three segments with low incidence of sideswipes was 18,000 for segment 14.0, 42,700 for segment 5.4, and 49,000 for segment 57.1. The median maximum AADT per lane for all 38 segments was approximately 54,300.

Rear-end collisions. Rear-end collisions represented relatively high percentages of all truck-involved accidents on segment 110.3 and intersecting segments 405.3 and 101.2 (Table 11). These are three of the heaviest traveled freeway segments in the area. In contrast, rear-end collisions represented relatively low percentages of accidents on segments 101.4 and 710.2, segments that are much less traveled.

The percentage of accidents that were rear-end collisions was found to be significantly related to the mean average annual daily traffic (AADT) at all locations along a freeway segment. The correlation between these two variables was 0.39 for the 38 segments ($p < .01$). Two of the three freeway segments with high incidences of rear-end collisions had the highest levels of mean AADT among all segments (206,300 for the segment 405.3 and 198,200 for the segment 101.2); the third segment (110.3) also had a high mean AADT level of 163,800. Correlations with maximum AADT and maximum AADT per lane were not significant. Thus, relatively high percentages of rear-end collisions are associated with higher levels of overall traffic, whereas high percentages of sideswipe collisions are associated with high levels of traffic per lane at key locations.

TABLE 10
ADJUSTED RESIDUALS FOR FREEWAY SEGMENT/COLLISION TYPE COMBINATIONS WITH SIGNIFICANT CELL EFFECTS IN THE LOG-LINEAR MODEL

	Collision Type					
Route	Sideswipe	Rear-end	Broadside	Hit Object	Overturn	Other Types
5.1				+2.3		
5.2					-2.8	
5.3	+2.6					
5.4	-5.8			+3.9	+4.5	
10.1						
10.2	+2.8			-2.2		
10.3					+2.2	
10.4						-2.4
10.5						
14.0	-4.5				+4.4	+4.2
22.0						
55.0						
57.1	-2.2		+2.1			
57.2						
60.1						
60.2						
91.1						+2.2
91.2						
101.1				-2.5		-3.5
101.2		+2.5				
101.3						
101.4		-2.3		+2.5		+3.0
110.1						
110.2				-2.6		
110.3		+2.8				
118.0			+2.6			
134.0						
210.1						
210.2						
405.1						
405.2			-2.5	-2.4		
405.3		+3.4	-2.3			
405.4						
605.1	-2.3					
605.2						
605.3				+2.3		
710.1			-2.2	+4.5		
710.2		-2.2				

Broadside collisions. With regard to broadside collisions, two segments had significantly high concentrations and three segments had significantly low concentrations. Segments 118.0 and 57.1 were high; and two adjacent segments of Route 405, 405.2 and 405.3, and segment 710.1 were low in terms of broadside collisions (Table 11).

Such collisions frequently occur on ramps (Table 4), and an investigation of the characteristics of the ramps for each freeway segment revealed that the percentage of broadside collisions was directly related to the percentage of ramps that are parts of diamond interchanges. Approximately 38% of all ramps in the study area on which truck-

TABLE 11
FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY HIGHER OR LOWER THAN EXPECTED: ALL COLLISION TYPES

Collision Type	Overall Average % of All Collisions	Significantly High Concentrations		Significantly Low Concentrations	
		Segment	% of All Collisions	Segment	% of All Collisions
Sideswipe	43.2	10.2	54.1	14.0	13.0
		5.3	49.6	5.4	25.4
				605.1	25.4
				57.1	33.3
Rear-end	31.3	110.3	46.4	710.2	15.6
		405.3	45.2	101.4	22.9
		101.2	45.1		
Broadside	4.8	118.0	11.8	405.3	1.1
		57.1	8.0	405.2	2.3
				710.1	2.9
Hit object	11.5	605.3	18.3	110.2	6.0
		5.4	17.6	101.1	6.8
		710.1	17.3	10.2	7.2
		101.4	16.9	405.2	7.9
		5.1	16.0		
Overturn	2.8	14.0	10.2	5.2	1.4
		10.3	7.8		
		5.4	6.2		
All other types	6.5	14.0	17.3	101.1	1.5
		101.4	11.4	10.5	1.8
		91.1	8.8		

involved accidents occurred are diamond-interchange ramps, but 77% of the ramps on segment 118.0 and 64% of the ramps on the segment 57.1 *are* diamond-interchange ramps. Conversely, only 26%, 10%, and 16% of the ramps on the three freeway segments with significantly low proportions of broadside collisions were diamond-interchange ramps. However, there are other freeway segments with very low or high percentages of diamond-interchange ramps, and further research is required to identify other possible causes of these collisions.

Hit-object collisions. High concentrations of hit-object collisions were found on segments 710.1, 5.4, 101.4, 605.3, and 5.1 (Table 11). Low concentrations were found on intersecting segments 405.2, 110.2, 10.2, and 101.1. The percentage of hit-object collisions was found to be the opposite of the percentage of rear-end collisions in terms of a relationship to traffic volume. The correlation between AADT at all points along a freeway segment and the percentage of hit-object collisions was $-.60$ ($p < .001$). It can thus be inferred that the number of hit-object collisions on an urban freeway segment does not vary as much with traffic levels as do the numbers of rear-end and sideswipe collisions. Consequently, the percentage of hit-object collisions is an inverse function of traffic levels whereas rear-end and sideswipe collisions are direct functions.

Overturn accidents. Segments with significant concentrations of overturn accidents are listed in Table 11. Segments 14.0, 5.4, and 10.3 had a high concentration, and segment 5.2 had a significantly lower percentage of overturns. Two of the three segments with high percentages of overturns, segments 14.0 and 5.4, are located in mountainous and rolling terrain. The third segment, segment 10.3, is adjacent to downtown Los Angeles and is built primarily with roadways on separate structures with relatively steep ramps. Further investigation is required to isolate other potential causal effects of such types of truck-involved accidents.

Other collisions. Finally, high percentages of "other" types of collisions were found on segments 14.0, 101.4, and 91.1, and low percentages were found on segments 10.5 and 101.1 (Table 11). As in the case of hit-object collisions, there is an inverse relationship between the percentage of other types of collisions and average traffic volume on a segment, with a correlation coefficient of −.65. However, the high incidence of other types of collisions on segment 91.1 demonstrates that other factors are involved as well, because the segment 91.1 has a greater than median level of AADT.

Ramp vs. highway accidents. Freeway segments with significantly higher or lower proportions of ramp accidents are listed in Table 12. Segments with relatively *high* concentrations of ramp accidents are intersecting segments 10.3 and 710.2, 605.3 and 10.5, 22.0 and 405.1, and segments 10.1 and 57.1, the majority of which are east and south of downtown Los Angeles. Segments with relatively *low* concentrations of ramp accidents (or high concentrations of highway accidents) are 101.1 and 101.3, 60.1, 5.3 and 5.4, 110.2, and 405.3, all of which are west or north of downtown Los Angeles.

On-ramp vs. off-ramp accidents. Three freeway segments were found to have relatively high concentrations of on-ramp (freeway entrance ramp) versus off-ramp (freeway exit ramp) accidents (Table 13). The overall split was 36% on-ramp, 61% off-ramp, and 3% other (such as truck scales and rest areas). However, these three segments, 605.2, 5.3, and 405.2, had from 50% to 63% on-ramp accidents. In contrast, segment 101.3 had fewer than 10% on-ramp (over 90% off-ramp) accidents.

TABLE 12
FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY HIGHER OR LOWER THAN EXPECTED: RAMP ACCIDENTS

Collision Type	Overall Average % of All Collisions	Significantly High Concentrations		Significantly Low Concentrations	
		Segment	% of All Collisions	Segment	% of All Collisions
Ramp accidents	16.8	10.3	41.2	101.1	8.4
		710.2	34.4	101.3	9.2
		22.0	31.0	60.1	9.3
		605.3	29.0	405.3	9.6
		57.1	26.3	5.4	9.7
		10.1	25.9	110.2	10.8
		405.1	23.5	5.3	12.2
		10.5	22.6		

TABLE 13
FREEWAY SEGMENTS WITH SIGNIFICANTLY HIGHER CONCENTRATIONS OF TRUCK-INVOLVED RAMP ACCIDENTS THAN EXPECTED: ON-RAMP VERSUS OFF-RAMP ACCIDENTS

Collision Location	Overall Average % of All Collisions	Segment	% of All Collisions
On-ramp	36.0	605.2	63.1
		5.3	51.2
		405.2	50.0
Off-ramp	61.0	101.3	90.9

Locations of ramp accidents. Three locations were distinguished for ramp accidents: ramp entry, ramp itself, and intersecting street (Table 14). Relatively high percentages of accidents at ramp entries occurred on two segments, 10.4 and 10.2, both of which serve the immediate downtown Los Angeles area. One segment, 710.1, had a high percentage of accidents on ramps. Finally, four segments had high concentrations of accidents on intersecting streets; these were 91.2 and 91.1, 5.1, and 57.1, all of which are at least partially in Orange County in the southern portion of the metropolitan area.

TABLE 14
FREEWAY SEGMENTS WITH SIGNIFICANTLY HIGHER CONCENTRATIONS OF TRUCK-INVOLVED RAMP ACCIDENTS THAN EXPECTED: THREE RAMP LOCATIONS

Collision Location	Overall Average % of All Collisions	Segment	% of All Collisions
At ramp entry	14.4	10.4	37.5
		10.2	28.3
On ramp itself	32.0	710.1	54.3
On intersecting streets	25.7	91.2	70.0
		5.1	53.3
		57.1	42.9
		91.1	40.8

Time period. The final accident characteristic investigated by freeway segment was the time period during which an accident occurred. Seven time periods were distinguished, during five of which there were significant differences in accidents on some freeway segments (Table 15). Three adjacent segments northwest of downtown Los Angeles had relatively high concentrations of accidents in the early morning hours (midnight to 6:00 a.m.). These were segments 101.1, 5.4, and 14.0, all of which are major truck routes north from Los Angeles. The 5.4 segment also exhibited a high percentage of accidents in the 9:00 p.m. to midnight period. Two segments, 57.1 and

TABLE 15
FREEWAY SEGMENTS WITH SIGNIFICANTLY HIGHER CONCENTRATIONS OF TRUCK-INVOLVED RAMP ACCIDENTS THAN EXPECTED: TIME OF DAY

Time of Day	Overall Average % of All Collisions	Segment	% of All Collisions
Midnight - 5:59 a.m.	7.0	14.0	26.5
		5.4	19.4
		101.1	10.8
6:00 a.m. - 8:59 a.m.	17.2	57.1	25.8
		10.5	22.5
9:00 a.m. - 11:59 a.m.	21.4	405.3	30.9
		5.2	25.0
noon - 2:59 p.m.	22.4	110.1	42.4
		110.3	32.7
		110.2	31.2
9:00 p.m. - 11:59 p.m.	4.6	5.4	14.4

10.5, had high percentages of accidents during the morning peak hours. Two segments, 405.3 and 5.2, had high percentages in the 9:00 a.m. to noon period. Finally, the three segments making up the entire length of Route 110 (110.1, 110.2, and 110.3) exhibited high concentrations of accidents in the noon to 3:00 p.m. period; Route 110 is the major harbor access route. No segments had significantly high or low concentrations of accidents during the afternoon peak hours or during the 6:00 p.m. to 9:00 p.m. period.

CONCLUSIONS

The method of log-linear modeling is potentially a powerful tool in identifying factors that underlie the relative frequency of occurrence of various accident characteristics, such as accident type, location, and severity. In the absence of interaction effects, the method can be used to obtain a direct measure of the degree to which any accident characteristic represented in a standard contingency table varies from its expected value.

In the application presented in this paper, the method was used to associate accident characteristics with type of collision and to identify freeway segments on which various accident categories were more prevalent than expected. The results indicated substantial differences between the types of collisions that tend to occur at ramp locations and those that occur along the freeway. The analysis was also able to uncover significant differences among the factors associated with the types of collision and to associate other characteristics, such as weather and road conditions, with particular collision types. In the analysis of accident characteristics by freeway segment, the analysis revealed several freeway segments that were particularly susceptible to certain types of accidents.

Some roadway characteristics, particularly overall traffic levels, were found to explain the pattern of freeway-segment results. However, the present research stops short of a thorough investigation of potential causal factors that distinguish freeway segments in terms of the types of collisions that occur on them.

Directions for further research include a multivariate analysis to relate the identified freeway-segment accident characteristics to roadway characteristics such as geometric design, shoulder provisions, and traffic patterns. The log-linear model residuals calculated in the present research could be used directly in future analysis. It would also be useful to contrast truck-involved accident characteristics with car-only accident characteristics as a control sample. The overall objectives of such future research would be to enhance truck-related safety considerations in freeway design.

REFERENCES

Baker, R. J., & Nelder, J. A. (1978). *The GLIM system. (Release 3).* Oxford: Royal Statistical Society, Numerical Algorithms Group.

Birch, M. W. (1963). Maximum likelihood in three-way contingency tables. *Journal of the Royal Statistical Society (Series B), 25,* 220–233.

Bishop, Y. M. M., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis.* Cambridge, MA: M.I.T. Press.

Bock, R. D., & Yates, G. (1973). *MULTIQUAL: Log-linear analysis of nominal or ordinal qualitative data by the method of maximum likelihood.* Chicago: International Education Services.

Caltrans. (1978). *Manual of traffic accident surveillance and analysis system.* Sacramento: State of California, Business and Transportation Agency, Department of Transportation, Division of Operations, Office of Traffic Engineering.

Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics, 10,* 417–451.

Golob, T. F., Recker, W. W., & Leonard, J. D. (1987). An analysis of the severity and incident duration of truck-involved freeway accidents. *Accident Analysis and Prevention* (forthcoming).

Goodman, L. A. (1972). Some multiplicative models for the analysis of cross-classification data. In L. LeCam, et al. (Eds.), *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. I, pp. 649–496) Berkeley: University of California Press.

Goodman, L. A. (1978). *Analyzing qualitative/categorical data.* London: Addison Wesley.

Haberman, S. J. (1973a). Log-linear models for frequency data: Sufficient statistics and likelihood equations. *Annals of Statistics, 1:* 617–632.

Haberman, S. J. (1973b). The analysis of residuals in cross-classification tables. *Biometrics, 29,* 205–220.

Haberman, S. J. (1974). *The analysis of frequency data.* Chicago: University of Chicago Press.

Haberman, S. J. (1978). *Analysis of qualitative data* (2 Vols). New York: Academic Press.

McCullagh, P., & Nelder, J. A. (1983). *Generalized linear models*. London: Chapman and Hall.

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalised linear models. *Journal of the Royal Statistical Society (Series A)*, *135*, 370–384.

Plackett, R. L. (1962). A note on interactions in contingency tables. *Journal of the Royal Statistical Society (Series B)*, *24*, 162–166.

Plackett, R. L. (1974). *The analysis of categorical data*. London: Griffin.