71-27,598

BEHNAM, Jahanbakhsh, 1935-
MULTIVARIATE ANALYSES OF TRAFFIC ACCIDENTS AT BRIDGE APPROACHES AND THEIR RELATIONSHIPS TO GEOMETRICS OF ROADWAYS.

The University of Oklahoma, Ph.D., 1971
Engineering, civil

THE UNIVERSITY OF OKLAHOMA

THE GRADUATE COLLEGE

MULTIVARIATE ANALYSES OF TRAFFIC ACCIDENTS AT BRIDGE APPROACHES AND THEIR RELATIONSHIPS TO GEOMETRICS OF ROADWAYS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

BY

JAHANBAKHSH BEHNAM

Norman, Oklahoma

1970

MULTIVARIATE ANALYSES OF TRAFFIC ACCIDENTS
AT BRIDGE APPROACHES AND THEIR RELATIONSHIPS
TO GEOMETRICS OF ROADWAYS

APPROVED BY

DISSERTATION COMMITTEE

ACKNOWLEDGEMENTS

The successful completion of this research was made possible through the cooperation and assistance of many individuals at the School of Civil Engineering and Environmental Sciences, and the Computing Center of the University of Oklahoma, and the Oklahoma Department of Highways. The author appreciates this opportunity.

In particular, the author is deeply grateful to the assistance and encouragements given him by Dr. Joakim G. Laguros, Professor of Civil Engineering, during the course of this research and throughout the graduate work of the author at the University of Oklahoma. Appreciation is expressed to Dr. John C. Brixey, Professor of Mathematics, for his assistance in discussing the mathematical aspects of the study. The author would like to thank Dr. Richard D. Bauman (formerly Assistant Professor at the University of Oklahoma) for his critical review of the manuscript and helpful suggestions during the early stages of the research. A special appreciation is extended to Professors Joe W. Keeley and J.

Lee Rodgers for kindly serving on the advisory committee and reading the manuscript.

The author is indebted to Doctors Roy B. Deal and Donald Parker, Biostatistics Department, the University of Oklahoma Medical Center, for discussing the mathematical modeling of the study. Dr. Deal's suggestions have been extremely helpful in developing the models.

Gratitude is also expressed to many individuals at the Oklahoma Department of Highways, particularly to Messrs. C. Dwight Hixon, Chief Traffic Engineer, and James W. Sparks, Traffic Planning Engineer, for their kind cooperation in providing traffic accident data. Also, Mr. Joe Kyle's assistance in collecting the data is appreciated.

A special note of appreciation is expressed to Mr. Kenneth W. Bray, the University of Oklahoma Computing Center, for his generous assistance throughout the computer analysis of this study.

The author is deeply indebted to a very special individual, his wife Joyce, for her editorial comments, review of manuscript, and many helpful suggestions offered during the preparation of the manuscript. Her devotion and encouragement throughout the entire graduate work of the author is appreciated.

Jahanbakhsh Behnam

Norman, Oklahoma
December 14, 1970

Dedicated to the memory

of my late father,

LUTF-ALLAH BEHNAM

TABLE OF CONTENTS

LIST OF TABLES

Table		Page

Table | Page

LIST OF ILLUSTRATIONS

MULTIVARIATE ANALYSES OF TRAFFIC ACCIDENTS
AT BRIDGE APPROACHES AND THEIR RELATIONSHIPS
TO GEOMETRICS OF ROADWAYS

CHAPTER I

INTRODUCTION

General

In the past two decades, evolution in highway transportation and the automobile industry has greatly influenced the nature of the American economy and its societal goals. Although this change and progress have brought increased prosperity to the nation's economy, they have also created tragic events of highway traffic accidents which annually embrace a staggering economic loss to the society. A recent report (1)* of the Department of Transportation revealed that a total of 513,000 fatal or serious injury type of accidents occurred on the highway system in the

* The number in parenthesis refers to references on page 175.

United States during 1969. The fatal accidents counted 59,000 of the total, while injury type accidents were estimated at 454,000. These accidents, not considering the minor injuries and "fender-bender" type of accidents, have cost the American society more than nine billion dollars in terms of lost services and damages incurred. These figures, of course, do not reflect unmeasurable costs such as human misery and suffering resulting from these car crashes.

These statistical data are not to provide the reader with illusive figures, but rather to point out the importance of traffic safety as one of the primary problems of the society today. Unfortunately, no coordinated or well-thought-out treatment has been taken toward effectively correcting the national problem of this scale. This lack of a productive program was documented in the President's safety message to the Congress in 1966 (2):

> Weaknesses of our present highway safety program:
> --Our knowledge of causes is grossly inadequate. Expert opinion is frequently contradictory and confusing.
> --Existing safety programs are widely dispersed. Government and private efforts proceed separately without effective coordination.
> --There are no clear assignments of responsibility at the federal level.
> --The allocation of our resources to highway safety is inadequate.
> --Neither private industry nor government officials concerned with automative transportation have made safety first among their priorities. Yet we know that

> expensive freeways, powerful engines and smooth exteriors will not stop the massacre on our roads.

It is clear that the 1966 Safety Act has improved the situation to some extent, and it has also increased the involvement of the Federal Government in the national problem of traffic accident prevention. The important fact remains that the process of implementing the Safety Act has been quite slow and the lack of participation and failure still exist at lower levels of government. Furthermore, little effort has been directed toward competent collective research in outlining the causes of traffic accidents with the exception of a few isolated small studies. The importance of research can not be emphasized further because a danger exists when opinion without scientific support becomes the sole basis for change. This type of change, if implemented, can cause damages in lives lost and money wasted (3).

Although the specific causes of traffic accidents are not known clearly, the general philosophy of highway accident causation is attributed to three basic components: (a) human factors, (b) vehicle factors, and (c) environmental factors. The detailed classification of the basic components and their present state of knowledge are shown in Table 1. It is apparent that the reduction in traffic fatality and injury problems can be achieved only through

TABLE 1

STATE OF KNOWLEDGE AND NEED IN TRAFFIC SAFETY AREAS

Human Factors	Environmental Factors	Vehicular Factors	State of Knowledge	Importance
Biographical Factors			Good	N/R[a]
Driving as a Skill			Poor	N/R
Medical Factors:				
Diseases			Poor	Minor
Physiological Impairments			Poor	Minor
Drugs and Chemicals			Poor	Minor
Alcohol			Fair	Critical
Personality Factors			Poor	Moderate
Driver Education			Poor	Moderate
Pedestrians			Poor	N/R
	Physical Factors:			
	Roadway		Good	Moderate
	Roadside		Fair	Major
	Road Discontinuities		Poor	Major
	Informational Factors:			
	Visibility		Poor	Major
	Communication and Signaling		Poor	Moderate
	Operational Traffic Control Factors		Poor	Moderate
		Sensory Factors	Poor	Moderate
		Layout Factors	Speculative	Minor
		Dynamic Control Factors	Poor	Moderate
		Vehicle Condition Factors	Poor	Minor
		Brakes and Tires	Poor	Moderate

[a] N/R: Not Rated

Source: The State of the Art of Traffic Safety, by Arthur D. Little, Praeger Publishers, New York, 1970, pp. 1-127.

collective study of these components as shown in Table 1. For example, it is known that the driver is the prime controller of his vehicle and he is required to make complex decisions under a variety of conditions. One may immediately conclude that the driver is a prime contributor to accident occurrence. It is clear, of course, that the driver is one of the fundamental factors contributing to the traffic safety problem, but can a driver be expected to operate his vehicle error-free at all times and under all conditions? Since the driving task requires a performance of complex perceptual-motor skills with the driver interacting with a large set of stimuli, the behavior and likelihood of error of the driver can be modified by changing the stimuli which may create conditions that produce driver fatigue, visual errors, and a consequent deterioration in performance. Based on this type of logic, the driver may be no more a contributor to accidents than the vehicle or roadway (4). Despite this, little scientific data regarding the driver's behavior are available, and the state of this knowledge is preliminary, as was indicated by Moynihan (5) in his report to the Secretary's Advisory Committee on Traffic Safety:

> Unfortunately, the present state of knowledge as to effectiveness of driver education provides no certainty, and much doubt, that the return on this enormous prospective effort will be commensurate with the

> investment. A broad and systematic inquiry is needed into the general question of how driving behavior is acquired, and how drivers can be taught not only to operate automobiles, but also to understand the major problems of highway safety.

This clearly states that there is a need for comprehensive research and an evaluation of driver behavior. To achieve a valid end-result, however, the diversion from conventional methodology is essential, and the emphasis should be placed on the scientific and intellectual methods (5).

It would be equally impractical to design very expensive vehicles equipped with highly sophisticated safety features without any improvement in human and environmental factors. Since the design of a safe automobile will require a high cost, the demand for such vehicles is dependent on the public decision, and it is very likely to be rejected. Although it is not a simple task to design automobiles to be safe for every driver, a high degree of direct control exists over the design of vehicles. Unfortunately, the relationships of vehicle design, accident, and other related factors are poorly understood at the present time. Safety research is urgently needed in vehicle design areas where man-machine elements can be studied so that their relationship to accidents can be established.

The environmental factors refer to the elements of roadway and surroundings which influence the initiation and severity of traffic accidents. The extensive improvement in roadway design combined with the construction of four-lane roadways temporarily reduced the frequency of fatal accidents between 1951 and 1964. For example, the traffic accident rate (fatal) declined from a post war high of 24.3 (per 100,000 population) to a low of 20.8 in 1961, and this rate increased to 24.9 in 1964 (4). It is quite clear that improvement in the environmental factors alone will not completely eliminate the traffic accident, but accidents can be reduced to a certain level of minimum. In contrast to two other factors (human and vehicle) discussed above, a great deal of research has been performed in roadway design and its relationship to accidents, however, the coverage has been sporadic and more definitive knowledge is still needed in many areas of design and traffic operation.

It is evident that the traffic accident will not be completely eliminated unless a collective approach is taken toward traffic safety. This will still leave responsibilities on traffic and highway engineers to continue their efforts in adding safety to newly designed highways and eliminating deficiencies on the existing roadways as much as possible.

Statement of the Problem

The concept of efficient and economical movement of people and goods has overcast the importance and the need for highway safety. In other words, in the past decade, the main objective of roadway designers and traffic engineers was directed toward the capacity and minimum delay concept while the safety aspects of the roadway were forgotten. The lack of high priority for safety is partially due to public attitude which Winfrey (6) summarized as:

> The American driving public does not want safe highways--that is, the people do not want to pay the financial cost and to suffer the restrictions necessary to produce safety in traffic. They would rather pay more for the accident costs associated with unsafe driving than to sacrifice their freedom to drive as they please.

The only time that the public becomes aware of a deficiency on a section of roadway and demands correction is when the portion of the roadway is the scene of numerous traffic accidents.

The cost of correcting deficiencies by spot improvement is high and it sometimes exceeds the original cost of construction (such as highway bridge structures, curve and grade removal). If no action is taken, the resulting economic losses from traffic accidents combined with human suffering become intolerable. Therefore, traffic and highway engineers must be provided with reliable evidence of relationships of traffic accident and geometrics of roadway. Only through such scientific facts the deficiencies on the existing roadways can be corrected intelligently.

The lack of adequate information regarding the relationship between traffic accident and geometrics of roadway at bridge approaches prompted the undertaking of this study. The main purpose of this research is to develop a predictive model based on geometrics of roadway at bridge approaches as parameters such that traffic accidents can be estimated through the use of this model. This type of model is extremely useful in studying the effects of each parameter on accidents. Utilization of this model will also permit identification of those parameters that contribute a great deal to accident occurrence at bridge locations. The construction of predictive models, where many parameters are to be included, requires application of a statistical method, namely, multiple linear regression analysis. Furthermore, attention is directed toward developing a model simple and practical as well as general in its application. The general format of the model is expected to be:

$$Y = a_o + \sum_{i=1}^{n} X_i$$

Since the characteristics of four-lane facilities are distinctly different from two-lane roads, this study will provide separate models for two-lane and four-lane highways. The application of the collected data will determine the specific format of the model for each classification of the roadway. Furthermore, the analysis of day versus night time accident, and dry versus wet conditions will be performed during the course of this research.

Significance of The Study

The limitations in economy and technology of roadway construction have prevented the designers from using optimum standards in the design of highways for safe operation. As the consequence of these limitations, the inadequate design of highways has resulted in sharp curvatures, steep grades, short sight distances, narrow roadways and structures, poor intersection design, and, in general, locating fixed objects near the pavement where the motorists can easily strike them. These spot deficiencies prompted the Federal Government to take a positive action in March 1964 when the President ordered the Bureau of Public Roads (presently Federal Highway Administration), an administrative agency of the Department of Transportation, to develop a program urging the local governments to launch action for removal of highway hazards (7). Furthermore, the examination of literature indicates that the Federal Government allocated $275,327,000 to spot improvement projects as of September 30, 1967, and after the states' participation, this figure totaled $758,206,000 (7). These figures simply show the necessity and importance of the spot improvement philosophy and its role in the reduction of accidents.

Accident occurrence at bridge approaches is not a new problem in highway safety. It has long been recognized that sudden discontinuity in roadway elements (pavement transition, sharp curve) create a situation with which the motorists are unable to cope. Location of a narrow bridge where steep grades lead to it or the existence of reverse curvature just before and after, characterize such discontinuities. For example, these types of bridges are frequently seen in the highway system in the State of Oklahoma. An inspection of the collected data indicates that there are numerous bridges with width less than 22.0 feet, and some randomly selected sites indicated bridge widths as low as 16.0 feet. Several examples of such bridge locations are shown in Figures 1 and 2. As indicated in the photographs, the grade and curvatures combined with a narrow bridge width create unsafe circumstances at high speeds. The examination

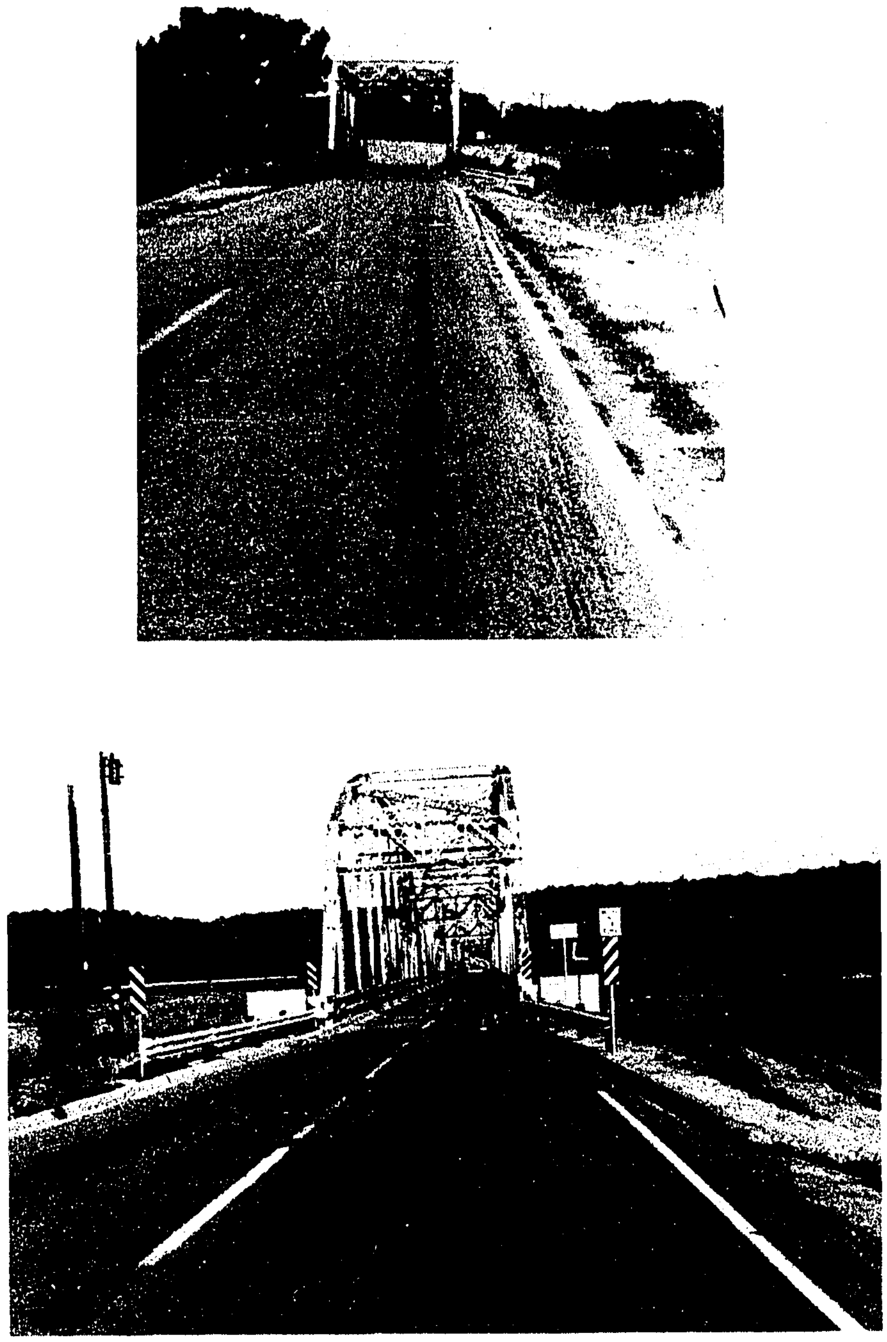

Figure 1. Selected Examples of Bridge Approaches, State of Oklahoma

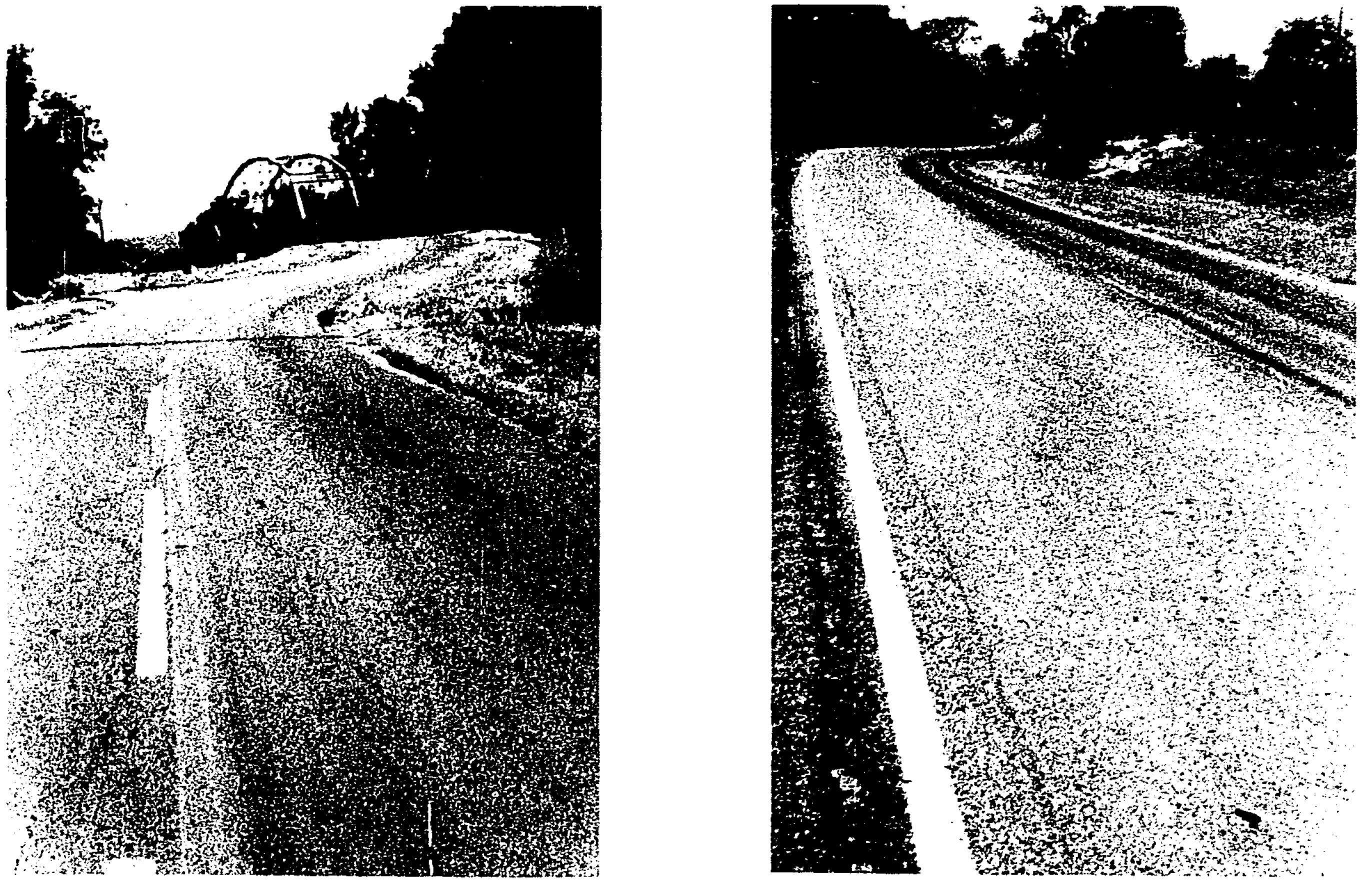

Figure 2. Selected Examples of Bridge Approaches,
State of Oklahoma

of accident records (8,9) reveals a higher property damage at bridge locations as compared to other fixed type objects of this nature. The bridge traffic accidents on Oklahoma Highways constituted 1,195 and 1,558 accidents for the years of 1968 and 1969, respectively. These clearly indicate that the bridge accidents increased approximately thirty-one percent during 1969. The property damages incurred were calculated to be 1,369,175 and 1,727,672 dollars for the calendar years of 1968 and 1969, respectively. It is evident that a loss of this magnitude can construct and finance at least an average of several bridges per year. This type of loss is an unnecessary economic loss which can be prevented by engineering studies resulting in highway modification.

Consequently, the need for complete analysis of bridge accidents is urgent, and the determination of accident relationship to geometrics of highways at bridge approaches and identification of causative elements should be considered a high priority project. The results of this research will enable practicing engineers to gain insight to the causes of accidents and will provide guidelines for taking action if any deficiency exists in roadway elements. Furthermore, such a study can provide criteria by which a priority of construction program can be established.

CHAPTER II

REVIEW OF LITERATURE

Introduction

During the conduct of this research, numerous publications relating to the subject area of this study were assessed and the content of each article was reviewed. The appraisal revealed, in general, that numerous investigations in geometrics of roadway have been performed, but they concentrated, to a large degree, on treatment of accidents on tangent sections with very few parameters included in the analysis. In addition, no comprehensive research was found to indicate a study of accidents at bridge approaches where all parameters regarding roadway geometrics were incorporated in the investigation. The following sections provide a brief summary of the literature reviewed, and the conclusions arrived at concerning the effect of roadway design elements on accident occurrence.

Characteristics of Traffic

In the following sections a review of traffic characteristics that are closely related to this study is presented. The literature cited will be concerned with the relationship of accidents to average daily traffic volume, travel speed, and effect of access control which will be used as parameters in developing the model in later chapters.

Average Daily Traffic

One of the first attempts to study the relationship of accidents to average daily traffic (ADT) was reported by Vey in 1937 according to the Automotive Safety Foundation (10). Vey's study concerning two-lane highways in New Jersey, indicated that the number of accidents (per million vehicle-miles) increased with the rise in ADT up to 7,000, and accident rates decreased as the volume increased over 7,000 vehicles per day. Moskowitz's work, as a discussion of Belmont's paper, which is based on the data collected in California on two and four-lane roads, showed that the accident per mile increased with ADT with no break in trend (when capacity not reached); however, Moskowitz's data had the same characteristics as Vey's in that both data converted to similar units and plotted (10). Furthermore, the accident study of two-lane roads by Schoppert (11) and Head (12), using multiple correlation analysis, indicated that ADT was directly related to accidents, and this interrelationship increased with higher ADT. These studies also indicated that a strong relationship existed between volume of traffic and other elements of roadway. The factor analysis of accident data by Versace (13), utilizing Schoppert's data, verified the existing relationship between ADT and accidents on two-lane highways.

Similar researches have been conducted in the study of traffic volume and highway accidents on four-lane roadways. The Interstate accident study, by Raff (14) in 1953, indicated that the number of accidents on four-lane highways was also directly related to volume of traffic. Two other studies by the Federal Highway Administration (15) and the Cornell Aeronautical Laboratory (16) concluded that traffic accidents increased as the volume tended to increase. The results of the Cornell Study revealed that the multi-vehicle traffic accident rate increased with the increase in ADT, but single car accident rates decreased at a higher volume of traffic.

Speed

The relationship of speed and accidents has been investigated by many researchers. According to the Automotive Safety Foundation (10), the most comprehensive study was accomplished by the Department of Commerce in 1959. The results of the study showed that on rural highways the accident rate increased at slow speeds and excessively high speeds, while it reduced to minimum at moderately high speed. The report further illustrated that, in general, accident rate was quite higher at speeds over 65 miles per hour and under 45 miles per hour. The report concluded that when the number of injury accidents is considered on miles-traveled basis, the minimum number of accidents occurs at the range of speed between 45-70 miles per hour.

Speed zoning and its effect on accidents was studied by several states throughout the United States, and the results showed a considerable reduction in accident rates after speed zoning was implemented (10). The studies in Michigan and California indicated 92 percent reduction in death rate and 70 percent reduction in injury type of involvement. Another study in Illinois showed interesting results; raising of the existing 30 miles per hour speed zones resulted in accident reduction of 20 percent where the 85th percentile speeds remained unchanged (10). A study of rural highways in Wisconsin by Mohr (17) also indicated reduction in accidents after speed zoning; however, in some cases fatal and injury accidents increased. Mohr's article states that if speed zoning were based on the traffic engineering studies and the resulting speed limits were properly posted, generally substantial reduction in frequency and severity of accidents should be expected.

The previously discussed studies simply considered the effect of speed zoning on accident rates by utilization of "Before and After" study techniques. It must be added that no scientific approaches were taken toward the evaluation of these changes.

Effect of Access Control

Access control of highways has a marked effect on the accident rates. A nationwide study by the Federal Highway Administration (10) concluded that the rate of accidents on fully controlled access highways was approximately one-half of those highways which were not access controlled. This reduction was also the result of other safety standards which were incorporated in the design of four-lane divided facilities. The study also indicated that rural accidents on controlled access facilities had higher severity than those in urban areas. Moskowitz's study (10) of the relationship of freeway accidents and average daily traffic volume pointed out that a certain portion of freeways had higher accidents than the established average accident rates and that deficiencies existed in the other features of freeway design where accident rates were higher than the established average.

A study on the Chicago expressways by Hoch (18) compared accident rates for expressways with the rate of accidents for arterials. The result of this study showed that expressways had 389 accidents per year fewer than arterial streets, based on 100,000 vehicles traveling both facilities for an equal distance. It is interesting to note that out of 389 accidents, 290 were classified as property damages, 98 were injury, and one was a fatality type of accident.

Cross Section of Roadway

Cross-section elements of highways have an important role in safety and traffic operations. Accident rates generally decrease with utilization of optimum standards in the design of roadway. The pertinent cross-section elements, namely pavement width, shoulder width, and width of structure, are reviewed in the following sections.

Width of Pavement

An early study, in 1932, by Morrison (10) regarding pavement width and safety, was oriented toward economic evaluation of pavement width in terms of saved accident costs. This study, which was performed on two-lane roadways in Michigan, showed that 20 feet wide pavements were safer than pavements of 15 or 18 feet in width. The research proved that savings resulting from accident reduction can justify the cost of widening rural two-lane roads from 18 feet to 20 feet.

Schoppert's (11) study of rural two-lane highways showed that, in general, the number of accidents increased when the cross-section (pavement and shoulder width) was decreased. In contrast, the study by Head (12) on urban extensions of state highways revealed that lane and pavement width did not have a strong correlation with accidents. Furthermore, he concluded that increase in the pavement width increased accident rates. This is an interesting conclusion because it shows that possibly the width of pavement in urban areas has an inverse relationship with accidents.

The AASHO Publication (19) also supports the relationship of narrow lane and pavement width on accident rates. In addition, it was pointed out that the capacity of narrow roadways is considerably lower than roadways with wider lanes and sufficient lateral clearance. The recommended standards by AASHO regarding roadway width are shown in Table 2.

The safety effects of pavement widening on two-lane roadways was also studied by Cope (10). The study consisted of widening 18 feet sections of roadway to 22 feet, and after widening, it was found that the reduction in accident rates ranged from 21.5 to 46.6 percent thereby indicating that the widening of roadway width became very significant in reducing accident rates at high traffic volumes. The results of the study are shown in Table 3.

Shoulder Width

Numerous articles concerning the relationship between shoulder width and traffic accidents have been published. Unfortunately, the studies were inconclusive and in some cases contradictory results were found. Belmont's study (20) showed that on a tangent and level road with a speed limit of 55 miles per hour, six feet shoulders were safer than narrow shoulders, and where the volume level was over 5000 vehicles per day, the shoulders that were wider than six feet were unsafe. Furthermore, another study (21) by Belmont indicated that when volume was over 2,000 vehicles per day, accident rates increased with paved shoulder width while this

TABLE 2

MINIMUM WIDTHS OF SURFACING FOR 2-LANE HIGHWAYS

Design Speed, mph	Current ADT 50-250	Current ADT 250-400	Current ADT 400-750 DHV 100-200	DHV 200-400	DHV 400 and Over
30	20	20	20	22	24
40	20	20	22	22	24
50	20	20	22	24	24
60	20	22	22	24	24
65	20	22	24	24	24
70	20	22	24	24	24
75	24	24	24	24	24
80	24	24	24	24	24

Source: A Policy on Geometric Design of Rural Highways, AASHO, 1965. (19)

TABLE 3

REDUCTION IN ACCIDENT RATES
RELATED TO
PAVEMENT WIDENING

Original Accident Rate	Percent of Reduction	Original Traffic Volume
Less than 1.5	21.5	2170
1.5 -- 1.9	25.2	2284
2.0 -- 2.4	34.4	2700
2.5 and up	46.6	3006

Source: Traffic Accident Experience--Before and After Pavement Widening, by A.J. Cope. Traffic Engineering, December 1955. (10)

rate decreased with volumes under 2,000 vehicles per day.

The study concerning rural two-lane highways by Head (10) showed that accident rates decreased when wider shoulder widths were used. Head, by grouping vehicle volume, was able to use the multiple correlation technique to establish a relationship between traffic accidents and widths of shoulders. Later, Head's findings were verified by Telford, Helwer, and Stohner (10).

In an effort to establish a relationship between accident rates and shoulder width, Billion and Stohner (22) classified the highways into several groups according to curvature, grade, and level tangent sections. The findings of this study revealed that the highways with wider shoulders had lower accident rates than those with narrow shoulders and it was concluded that the accident rates on roadway sections with curves and grades were higher than those on level tangent sections.

These studies simply pointed to the existence of a general relationship between shoulder width and accident rates on highways with similar characteristics but no specific relationship between the variation of shoulder width and traffic accident rates. However, an unpublished article (10) indicated that when the existing shoulders (one to four feet in width) were widened to six feet on a 20 feet roadway, the accident rate of 2.76 per million vehicle-miles was reduced to 1.86 per million vehicle-miles.

Width of Bridge Structure

Very little research has been conducted concerning the relationship between the width of bridge structure and traffic accidents. The earliest publication, in 1955 by Williams and Fritts (10), concerned a study of the influence of width of bridge structures on traffic accidents. It was concluded that a reduction in accidents occured when the bridge structure was from one to five feet wider than the approaching pavement. This decrease in accident rates is shown in Table 4.

Gunnerson (23) studied the effect of bridge width on accidents at 65 bridges which were widened to 24 feet and seven control bridges having 30 feet width on two-lane roadway with 24 feet width. The investigation consisted of "Before and After" study where the volume of traffic was grouped and accident types were discriminated. Tables 5 and 6 show the data used and the accident reduction resulting from this modification. Thus, Gunnerson concluded that when the roadway was widened without widening of the bridge structure, the accident rates increased; widening of bridge structure

TABLE 4

ACCIDENT RATE AS INFLUENCED BY
WIDTH OF STRUCTURE

Bridge Width Wider Than Width of Approaching Pavement	Accident Rate Per 100 Million Vehicle-Miles
Less than--1 foot	100
1--5 feet	58
5--Over	12

Source: Let's Build Safety into our Highways, by Sidney J. Williams and Carl E. Fritts, Public Safety, May 1955. (10)

TABLE 5

SUMMARY OF ACCIDENT AND BRIDGE DATA FOR NARROW BRIDGES IN IOWA
(1948 through 1959)

	Highway Surface Widened Only for Given Ranges of ADT						Highway Surface and Bridge Widened	
	0 -- 999		1,000 -- 2,999		3,000 -- Over		1,580 -- 3,630	
	Before	After	Before	After	Before	After	Before	After
Number of Bridges Studied	12	12	46	46	7	7	7	7
Bridge-Months	902	826	4,772	1,852	638	370	795	213
Total Accidents	5	25	144	95	23	32	28	1
Bridge Hit	1	13	61	52	11	14	7	1
Injuries	2	18	74	62	11	14	19	2
Property Damage, $	990	12,028	100,856	92,040	16,267	16,353	14,291	1,700

Source: Iowa Narrow Bridge Accident Study, by H.E. Gunnerson, Highway Research Abstracts, Volume 31, No. 7, 1961.

TABLE 6

SUMMARY OF ACCIDENT AND BRIDGE DATA FOR NARROW BRIDGES IN IOWA
(ACCIDENT DATA PER 1,000 MONTHS)

	Highway Surface Widened Only for Given Ranges of ADT						Highway Surface and Bridge Widened	
	0 -- 999		1,000 -- 2,999		3,000 -- Over		1,580 -- 3,630	
	Before	After	Before	After	Before	After	Before	After
Number of Bridges Studied	12	12	46	46	7	7	7	7
Accidents	5.5	30.3	30.2	51.3	36.0	86.5	35.2	4.7
Hit Bridges	1.1	15.7	12.8	28.1	17.2	37.8	8.8	4.7
Injuries	2.2	21.8	15.5	33.5	17.2	37.8	23.9	9.4
Property Damage, $	1,098	14,562	21,134	49,698	25,497	44,197	17,976	7,981

Source: Iowa Narrow Bridge Accident Study, by H. E. Gunnerson, Highway Research Abstracts, Volume 31, No. 7, 1961.

combined with widening of roadway width resulted in reduction of accident rates.

The two studies cited above indicate that relative width (difference between bridge width and approaching pavement width) has a significant effect on accident rates. For example, one should expect a 22 feet roadway with a 22 feet bridge structure to be safer than a 22 feet bridge on a 24 feet roadway.

Roadway Alignment

The influence of alignment on traffic accidents has been known for many years; this influence is particularly pronounced on rural two-way highways. Previous studies of roadway alignment have established a definite relationship between traffic accident and roadway alignment. Early researches regarding alignment (horizontal and vertical) and accident will be reviewed in the sections below.

Horizontal Alignment

Raff's (14) study of two-lane and four-lane roadways revealed that accident rates increased as the result of higher degrees of curvature. Accident rates were 0.15 and 0.40 per degree of curve on rural two-way and four-lane facilities, respectively. The study by Kipp (10) indicated similar results which are tabulated in Table 7. Billion and Stohner (22) and Giles (10) documented that sections of roadways with curvatures over five degrees were 19.27 to 23.0 times more accident prone than straight sections of highways.

In addition, the effect of frequency of curvature on accident rates has been studied on rural roadways. Raff (14) could not establish a significant relationship between accident and frequency of curvature, but the study by Baldwin (10) showed that accident rates decreased as the frequency of curvature increased. The Automotive Safety Foundation reports that Kipp (10), in studying effect of frequent curvature on accident, verified the conclusions reached by Baldwin. Kipp's study indicated that curves at the end of long tangent roadways were more hazardous than curves at the end of short tangent highways.

Thus, it is clear that curvature has definite effects on accident occurrence on two-lane and four-lane rural highways. Furthermore, the frequency of curvature decreases the accident rates on rural highways.

Vertical Alignment

Vertical alignment has been a major problem on two-lane rural highways, and its relationship to traffic accidents has been investigated by many authors. Some of

TABLE 7

ACCIDENT RATES RELATED TO CURVATURE

Degree of Curve	Accident Rate (per million vehicle-miles)
Less than 3	1.37
3 -- 5	2.48
5 and over	3.86

Source: Minnesota Roadside Survey: Progress Report on Accident, Access Point and Advertising Sign Study in Minnesota, by O. L. Kipp, Highway Research Board, Bulletin 38, 1951. (10)

the studies cited under "Horizontal Alignment" also included the effect of vertical alignment on accidents, and the conclusion revealed that a direct relationship existed between traffic accidents and grade (vertical alignment).

The Automotive Safety Foundation indicated (10) that Young studied the relationship of traffic accident and sight distances on two-lane roadways in California. His findings showed that traffic accident rates increased with insufficient length of sight distance available. This evidence is shown in Table 8.

Bitzel (10) studying expressways, found a marked relationship between gradient and accident rates. The results of this study are shown in Table 9. In a study of grades on freeways, Mullins and Keese (24) found that the majority of accidents occurred on crests and sags.

In another study on an Ohio Turnpike, Bowman (25) found that traffic accidents were related to grade. This was true even though upgrades on the turnpike were held to two percent and downgrades to approximately three percent. He further concluded that "reasonable downgrades experience accident rates that are slightly higher than those recorded for level sections," but that "upgrades even when held to two percent produce accidents." This study showed that

TABLE 8

ACCIDENT RATES RELATED TO SIGHT DISTANCES ON TWO-LANE ROADS

Sight Distance (feet)	Accident Rate (per million vehicle-miles)
Less than 800	2.4
800 - 1500	1.9
1500 - 2500	1.5
2500 - Over	1.1

Source: Building Safety into Our Road System, by J. C. Young, California Traffic Safety Conference, Proceedings, 1950. (10)

TABLE 9

ACCIDENT RATES RELATED TO GRADIENT

Gradient (percent)	Accident Rates (per 100 million vehicle/KM)
0 - 1.99	46.5
2 - 3.99	67.2
4 - 5.99	190.0
6 - 8.00	210.5

Source: Effect of Motorway Design on Accidents in Germany, by I. F. Bitzel, Highway and Bridges and Engineering Works, October 1956. (10)

downgrades produced fewer accidents than upgrades. The reason for this was partly attributed to heavy commercial vehicles which were climbing the freeway at much lower than normal speed.

A comprehensive study of vertical alignment by the Texas Transportation Institute (26) proposes that head-on collision criterion be used for determining design values of sight distance on two-lane roads. These design values, as shown in Table 10, were based on a 2.5 second perception-reaction time and the total distance needed to stop both opposing vehicles. The report states that although these values may seem liberal, these criteria consider the failure of the opposing vehicle to slow in case of a hazard where the other driver has to take evasive action in a very short time.

Woods (27) takes a different approach in computing "desirable" design criterion for sight distance. The study assumes zero height of object and standard height of eye (3.75 feet) in deriving the values of sight distance which are shown in Table 11. Woods, referring to the normal focal point of a driver's eye as derived by Matson (28), documents, "Although these figures would indicate a sight distance of 2,000 feet or more might be desirable, there is considerable doubt that the driver would be able to distinguish lane

TABLE 10

SAFE STOPPING SIGHT DISTANCES FOR STATIONARY OBJECT CRITERIA
(EMPLOYING ASSUMED SPEEDS DERIVED IN REPORT)

Design Speed	Assumed Speed for Conditions	Perception-Reaction Time	Perception-Reaction Distance	Friction Factor	Braking Distance On Level	Stopping Sight Distance Computed	Stopping Sight Distance Rounded
mph	mph	sec	feet	f	feet	feet	feet
30	34	2.5	125	0.28	138	263	250
40	42	2.5	154	0.26	226	380	380
50	50	2.5	184	0.24	347	531	530
60	59	2.5	217	0.23	504	721	720
65	62	3.0	273	0.23	557	830	830
70	64	3.0	282	0.23	593	875	880
75	67	3.5	345	0.22	680	1025	1030
80	70	3.5	360	0.22	742	1102	1100

Source: Evaluation of Stopping Sight Distance Design Criteria, by John C. Glennon, Texas Transportation Institute, Texas A & M University, College Station, 1970.

TABLE 11

DESIRABLE DESIGN CRITERIA

Design Speed (mph)	Height of Eye (feet)	Height of Object (feet)[a]	Sight Distance (feet)
50	3.75	0	600
60	3.75	0	750
70	3.75	0	900
80	3.75	0	1100

[a] Visual contact with the roadway surface provided.

Source: Vertical Alignment Effects on Safe Vehicle Operation, by Donald L. Woods, Texas Transportation Researchers, Texas A & M University, College Station, 1970.

markings or pavement edges at such a great distance. A design distance of this magnitude would probably be unrealistic." Referring to Table 11, he concludes that these values, as supplementary information, should be used with caution and in conjunction with AASHO standards for sight distance.

CHAPTER III

PROCESSING AND SOURCE OF DATA

Introduction

In order to gather the proper data, the Oklahoma State roadway system was divided into two classes namely two-lane and four-lane rural highways. The four-lane highways included the Interstate Highway System and other four-lane highways which had characteristics comparable to the Interstate Highways. The total intended sample was over 1500 distinct locations; however, lack of proper and valid data reduced this number to approximately 800 distinct locations. Finally, total locations with good data were classified--approximately 650 locations as two-lane rural and 150 locations as four-lane rural highways. The accidents analyzed occurred during 1968 and 1969 over the entire state of Oklahoma. The geographical distribution of these accidents by county is shown in Table 12.

In fulfilling the objectives of this study, it was necessary to collect detailed information on the geometrics

TABLE 12

GEOGRAPHICAL DISTRIBUTION OF BRIDGE ACCIDENTS IN OKLAHOMA

County	Population	1968 Accident		1969 Accident	
		Number	Property Damage, $	Number	Property Damage, $
Adair	14,646	5	1,925	3	2,100
Alfalfa	8,660	12	7,575	6	7,475
Atoka	12,124	16	18,125	15	13,950
Beaver	7,169	3	2,225	3	14,150
Beckham	18,298	8	7,200	9	6,275
Blaine	13,513	18	18,513	20	18,325
Bryan	27,145	20	28,850	16	31,350
Caddo	32,605	14	19,400	25	46,425
Canadian	28,330	8	20,775	10	34,950
Carter	38,555	13	19,150	7	8,400
Cherokee	21,340	1	550	7	4,225
Choctaw	17,236	11	10,625	4	2,375
Cimarron	4,662	4	3,525	2	550
Cleveland	79,913	10	6,275	22	21,500
Coal	6,284	3	17,875	4	5,850
Comanche	128,938	24	23,925	24	25,925
Cotton	8,383	17	61,575	12	9,200
Craig	16,562	8	4,708	9	8,800
Creek	44,455	23	27,375	36	31,075
Custer	24,102	9	6,450	3	4,200
Delware	14,335	1	600	1	750
Dewey	6,616	3	2,425	2	2,800
Ellis	5,976	5	7,025	4	3,675
Garfield	56,334	6	7,050	3	975
Garvin	28,681	12	9,700	18	15,125
Grady	32,209	17	17,900	29	45,775
Grant	8,291	1	825	8	8,700
Greer	9,660	5	1,825	3	18,125
Harmon	6,279	5	23,375	3	6,700

TABLE 12--Continued

County	Population	1968 Accident Number	1968 Accident Property Damage, $	1969 Accident Number	1969 Accident Property Damage, $
Harper	5,965	5	5,000	8	63,425
Haskell	10,148	8	7,250	11	31,625
Hughes	15,741	9	9,800	5	2,200
Jackson	30,774	11	33,525	12	14,300
Jefferson	8,590	3	435	8	12,025
Johnston	9,457	12	8,550	5	4,950
Kay	52,109	13	13,875	18	37,375
Kingfisher	11,945	4	2,875	5	9,325
Kiowa	16,613	11	13,075	9	15,600
Latimer	8,778	4	11,800	2	2,925
LeFlore	32,418	24	41,225	12	11,123
Lincoln	20,007	9	22,725	16	13,750
Logan	19,409	12	16,125	6	4,125
Love	6,387	12	17,025	4	1,750
McClain	13,983	3	1,100	11	11,875
McCurtain	30,352	4	4,875	9	20,400
McIntosh	14,573	4	1,375	10	7,850
Major	8,682	1	2,500	6	8,525
Marshall	7,865	3	12,350	1	2,550
Mayes	21,133	2	3,150	9	18,200
Murray	10,651	37	48,225	57	55,475
Muskogee	64,662	32	75,900	40	45,750
Noble	10,174	12	19,325	13	31,350
Nowata	10,863	2	1,225	7	9,925
Okfuskee	13,128	8	6,450	3	3,400
Oklahoma	477,663	297	195,725	514	387,649
Okmulgee	39,257	13	5,700	11	23,325
Osage	31,136	18	19,550	24	29,650
Ottawa	31,279	3	18,550	11	14,650

TABLE 12--Continued

County	Population	1968 Accident		1969 Accident	
		Number	Property Damage, $	Number	Property Damage, $
Pawnee	11,466	7	4,500	6	5,900
Payne	50,958	18	57,600	24	25,025
Pittsburg	39,158	11	22,800	11	11,450
Pontotoc	30,378	17	9,300	22	15,500
Pottawatomie	44,035	13	17,200	26	37,000
Pushmataha	10,148	4	7,575	6	6,100
Roger Mills	5,390	4	1,275	0	0
Rogers	22,875	12	14,975	20	21,875
Seminole	30,574	8	5,450	4	285
Sequoyah	20,058	22	21,250	16	29,650
Stephens	38,364	10	10,150	1	2,550
Texas	15,608	5	4,150	7	5,425
Tillman	15,294	10	9,300	12	12,925
Tulsa	369,347	151	99,175	195	169,150
Wagoner	17,554	20	36,125	23	33,675
Washington	42,964	2	975	2	5,000
Washita	19,861	16	47,700	7	5,975
Woods	13,427	2	250	7	14,850
Woodward	15,756	5	1,850	14	17,975
Total	2,568,258	1,195	1,369,175	1,558	1,727,672

of the locations involved and other pertinent data. A brief explanation of the data collection process on accident information, geometrics of roadway, and traffic data is described in the following sections.

Accident Information

The accident records which were utilized to obtain the necessary information for this study were made available by the Traffic Engineering Division of the Oklahoma Department of Highways. Comprehensive accident information was compiled on the computer tapes which provided the location, type, and much other pertinent data concerning the particular accident. The output listing consisted of ninety-four columns which grossly concentrated the information regarding condition and causes (driver behavior) from the viewpoint of the investigator. The characteristic of each accident is described by numerical numbers indicating the coding convention (29) which has been adopted to be compatible with the spacing problem and provide simplicity in processing. Unfortunately, except the width of pavement, this type of listing does not furnish any information regarding roadway elements. As a result, it was only possible to obtain the identification number (approximate location) and number of accidents from this listing. To accomplish the intent of this study, it was necessary to obtain the remaining data from elsewhere, as will be described later in this chapter.

Traffic Characteristics

The volume of traffic, in ADT, was obtained from several sources which are published annually by the Planning Division of the Oklahoma Department of Highways. The main source used to select ADT information was the Sufficiency Rating Report (30) which is also published by the Oklahoma Department of Highways. Although the variation of ADT within two years of the study period was not expected to be significant, the recorded volumes of traffic were cross checked with two other (31,32) publications to assure that the most reliable volumes were used in the analysis. The Department of Highways has two types of counting stations, the permanent with an inductive loop detector and the temporary manual, which cover the entire state highway system. The data collected by means of these counters (total of 146 counters) were reduced and tabulated into usable forms according to types and classifications.

The travel speeds were obtained from accident listings for each accident because this was the only information source from which such data could be gathered. The speed of a traveling vehicle before an accident refers to an estimate of traveling speed which is based on the evidences at the location of the accident at the time of investigation. Although this is a crude estimate of the speed, it enables one to consider the speed as a factor in the analysis. Furthermore, it was intended to include another variable as "speed change" (differential speed) in the analysis at bridge approaches. The investigation indicated that the quantitative values for the "speed change" variable could not be obtained from the sources available at the present time. Therefore, it was not included in the list of potential independent variables. Further treatment of speed will be given in later chapters.

Geometrics of Roadway and Bridge

The geometrics of roadway and bridge information provided the bulk of the data needed to conduct this study. The description of roadway elements considered in the analysis is given in Table 13.

The accident locations and number of accidents were determined from accident listings as described previously. Since the geometrics of roadway were not available from the

TABLE 13

DESCRIPTION OF GEOMETRIC ELEMENTS
OF
ROADWAY AND BRIDGE

Geometric Elements	Description
Width of Bridge	Measured Curb to Curb
Width of Approaching Pavement	Measured Edge to Edge
Sight Distance (Critical)	Measured as Prescribed by AASHO Standards (19).
Approaching Grade (Critical)	Measured Within Defined Limits[a].
Degree of Curvature	Design Value as Specified in Plans Within Defined Limits.
Height of Guard Rails	Design Height as Specified in Plans.
Length of Bridge	Design Length as Specified in Plans.
Number of Openings	It is number of Driveways or Roadways Intersecting Highway Under Study Within Defined Limits.

[a] Defined Limits will be Described in Chapter V.

same source, the Maintenance Section of the Bridge Division, the Oklahoma Department of Highways, was contacted in the search for this type of information. Although the main function of this section is not to keep the record of geometric details, each bridge location is filed according to control numbers and it has a complete plan-profile combined with other information. Most of the data listed in Table 13 were obtained with reasonable accuracy through the use of these files. Since there were still some geometrics unavailable through these files, the central file of the Department was contacted in completing the remaining information.

The procedures utilized in measuring the variables indicated in Table 13 are fully described in the AASHO's handbook, A Policy on Geometric Design of Rural Highways (19). The sight distances and approaching grades were determined for either side of the bridge location under study. The critical value of these measurements was considered in the analysis, and the sign convention for approaching grade was adopted from AASHO's publication.

Index of Sufficiency Rating

In order for the Oklahoma Department of Highways to establish means of estimating future needs and determining a priority program for construction, the sufficiency rating method is used. The concept of sufficiency rating is to establish an index number for determining the level of adequacy based on comparison of the existing facility to specified design standards. The adequacy of any roadway is dependent on two factors: (a) design features of the roadway, and (b) the condition. The design features and condition are defined as shown in Tables 14 and 15. Each element of the roadway and condition is assigned a relative point value which, when perfect, adds to 100 points. According to this procedure, the higher score indicates an adequate roadway whereas a lower score indicates deficiency in the roadway.

Although this index is based on a subjective comparison method and has a limited scientific support, it was used as one of the variables in this study. The value of each score at each particular bridge location was obtained from the Sufficiency Rating Report (30).

TABLE 14

INDEX OF SUFFICIENCY RATING -- ROADWAY ELEMENTS

Description	Score
Surface Width	16
Surface Type.	8
Shoulder Width and Type	6
Curvature	8
Grade	5
Stopping Sight Distance	8
Passing Opportunity	8
Hazards	6
Total (Roadway Rating)	65

Source: Sufficiency Rating Report and Need Study, Planning Division, Oklahoma Department of Highways, Oklahoma City, 1968.

TABLE 15

INDEX OF SUFFICIENCY RATING -- CONDITION

Description	Score
Foundation	14
Weaving Surface.	10
Drainage	7
Shoulders.	4
Total (Condition Rating).	35

Total Index of Sufficiency Rating:

65 + 35 = 100

Source: Sufficiency Rating Report and Need Study, Planning Division, Oklahoma Department of Highways, Oklahoma City, 1968.

CHAPTER IV

MATHEMATICAL CONCEPTS IN MODEL BUILDING PROCESS

Introduction

The objective of researching a mathematical approach to engineering problems is to base the ultimate solution on an exact science which can describe the events of the real world in abstract forms of equations. Unfortunately, in the field of engineering inability to make accurate measurements of parameters, inability to identify particular variables, and sometimes simply lack of knowledge, prevent researchers from establishing a mathematical relationship to describe certain phenomena. Since the concept of cause and effect is often obscure in scientific research, it is possible to use independent variables, which can be easily measured and identified, in predicting a certain event whose behavior cannot easily be investigated. In such a case, the linear statistical models are capable of finding a relationship that can predict the events of the real world.

The concept of linear statistical inference will be briefly reviewed in this chapter, and the essential aspects of the method relevant to this study as applied science will be discussed. Since an in-depth mathematical analysis of the subject can be found in the statistical publications and it is beyond the scope of this study, no mathematical proof of the material will be shown in this chapter.

Definitions

The following definitions will be used throughout the remaining chapters of this study:

Independent variable--a variable such that its values can be observed without any control or can be pre-selected.

Dependent variable (response variable)--a variable such that its values are determined as the result of changes in the values of independent variable(s).

Model--of a real phenomenon is a mathematical theory established by the axiomatic method with a rule for interpreting propositions of the mathematical theory into propositions concerning the real phenomenon (59).

Linear Model--is a mathematical expression, involving random variables, mathematical variables, and parameters, which is a linear relationship in the random variables and in the parameters (33).

Moment--expected values of the powers of a random variable which has a given distribution (34).

Mean--the first moment of a given distribution.

Variance--the second moment about the mean. The square root of variance is called standard deviation.

Coefficient of Correlation--if x and y are two random variables with a joint distribution, the ratio of their covariance to the square root of the product of their variances is called the coefficient of correlation (50).

Multivariate Statistical Analysis

The most well-known and common distribution in statistical analysis is normal distribution with one variable. In contrast, the multivariate normal distribution is a statistical density that consists of several normally distributed variables. The analysis of multivariate normal distribution involves complex and tedious processes as the number of variables increases. Fortunately, the utilization of regression models, namely the method of least squares (used when non-normality may exist in data), can accomplish the objectives of this study. Consequently, a linear regression model combined with the use of matrices will greatly simplify the task of illustrations in this chapter.

If a linear relationship exists between a dependent variable, y, and independent variables $x_1,x_2,\ldots,x_n$, and assuming that n observations were taken on p variables, a linear regression model can be expressed by the equation below:

$$y_i = a_1 + a_2x_{i2} + \ldots + a_px_{ip} + e_i \quad (i = 1,2,3,\ldots,n) \qquad (4.1)$$

In the equation above the a_i are parameters, the x_i are assignable mathematical variables, and e_i are random variables where: $E(e_i) = 0$, $V(e_i) = \sigma^2$, and $Cov(e_i,e_j) = 0$ for $i \neq j$, $i = 1,2,3,\ldots,n$, $j = 1,2,3,\ldots,n$. The equation in matrix form is shown below:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & x_{12} & \cdot & \cdot & \cdot & x_{1p} \\ 1 & x_{22} & \cdot & \cdot & \cdot & x_{2p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & x_{n2} & \cdot & \cdot & \cdot & x_{np} \end{bmatrix}, \quad A = \begin{bmatrix} a_1 \\ a_2 \\ \cdot \\ \cdot \\ \cdot \\ a_p \end{bmatrix}, \quad E = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_n \end{bmatrix}$$

Where:

Y: (nxl) Vector presenting observations on ys

X: (nxp) Matrix presenting observations on xs

A: (pxl) Vector presenting unknown parameters

E: (nxl) Vector presenting error terms or residuals

Equation (4.1) now can be written in a simpler form as:

$$Y = XA + E \qquad (4.2)$$

The assumptions regarding equation (4.2) are that $E(E) = 0$, $V(E) = I\sigma^2$ (unknown constant variance σ^2), and e_i terms have a normal distribution and are not correlated.

The method of least squares will be employed to estimate the parameters in equation (4.2). This method is based on the principles of minimizing the sum of the squares of deviations (dependent variable) of all points from the regression line in a sample under study (35). Since $E(E) = 0$ and $E(Y) = XA$, the sum of the squares of the error terms can be computed as follows:

$$E'E = (Y - XA)'(Y - XA) \tag{4.3}$$
$$= Y'Y - A'X'Y - Y'XA + A'X'XA$$

Since $A'X'Y$ is a scalar matrix, it follows that $(A'X'Y)'=Y'XA$. Then, substituting this equality in equation (4.3) and combining similar terms, the result is:

$$E'E = Y'Y - 2A'X'Y + A'X'XA \tag{4.4}$$

In order to minimize the quantity $E'E$ to obtain estimates of A, it is necessary to differentiate equation (4.4) with respect to A, equate to zero, and solve for A. The solution for A is possible when the matrix $(X'X)$ is a nonsingular matrix and has an inverse, and since the x_i are mathematical variables it is possible to choose them such that the matrix $(X'X)$ is nonsingular and has an inverse. The estimate of A is:

$$\hat{A} = (X'X)^{-1}X'Y \tag{4.5}$$

The estimate of A obtained by the method of least squares has the following desired properties (36):

1. $\hat{A}$ is an estimate of A which minimizes the sum of the squares of error terms, $E'E$.

2. $\hat{A}$ is an unbiased estimate of A and it has a minimum variance, $\mathrm{Var}(\hat{A}) = \sigma^{2}(X'X)^{-1}$, among all linear

unbiased estimates. An unbiased estimate refers to mean values of estimates that are equal to the population mean.

3. According to assumptions made previously, error terms are independent and approximately normal with a mean of zero and variance σ^2. This implies that $\hat{A}$ is the maximum likelihood estimate of A and indicating that E has the following properties:

a. E is distributed as a multivariate normal distribution.

b. E (E) = 0

c. V (E) = I σ^2, I being an identity matrix.

Analysis of Variance

Analysis of variance is a statistical testing procedure which has numerous applications in statistical investigations. This technique will be described in connection with its application in linear regression models. The analysis of variance can be used to investigate and determine the precision of predictability of the regression line.

Let Y_i, $\hat{Y}_i$, and $\bar{Y}$ denote the observed, estimated, and mean values of y, respectively. By definition the following relationship exists:

$$e_i = Y_i - \hat{Y}_i = Y_i - \bar{Y} - (\hat{Y}_i - \bar{Y}) \qquad (4.6)$$

If both sides of equation (4.6) are squared, summed, and necessary transformations are made, the following equation will result:

$$\sum(Yi - \bar{Y})^2 = \sum(Yi - \hat{Y}i)^2 + \sum(\hat{Y}i - \bar{Y})^2 \qquad (4.7)$$

Equation (4.7) is a basic and important relationship in the analysis of variance. The individual terms of the equation are described below:

$\sum(Y_i-\bar{Y})^2$: The sum of the squares of the deviations of the observations from the overall mean (commonly known as "SS about the mean").

$\sum(Y_i-\hat{Y}_i)^2$: The sum of the squares of the deviations of the observations from their predicted values (commonly known as "SS about regression").

$\sum(\hat{Y}_i-\bar{Y})^2$: The sum of the squares of the deviations of the predicted values from the overall mean (commonly known as "SS due to regression").

A study of equation (4.7) indicates that the variation in SS about the mean is dependent on the variations of SS about regression ("unexplained" SS) and SS due to regression ("explained" SS). Furthermore, if a regression line were fitted perfectly to the data available, the "unexplained" SS will reduce to zero indicating that the variation is due to regression only. In the analysis of the regression model, one should try to keep the variation of "unexplained" SS to a minimum in order that all variations can be shifted to the "explained" SS. This means that the ratio of SS due to regression to SS about the mean should be converged to unity as much as possible. Therefore, the analysis of variance is a means by which one can find the degree of variation that each term contributed to SS about the mean. In general terms by utilization of matrix notations, the analysis of variance (AOV) may be formed as shown in Table 16.

Multiple Discriminant Analysis

A general meaning of the word "discriminant" is defined by Webster as: "A mathematical expression providing a criterion for the behavior of another more complicated expression, relation, or set of relations." Although this definition describes the intent of discriminant functions, the application of this statistical method, as used in applied science, will be reviewed. The theory of discriminant functions which has numerous uses in the multivariate analysis was developed by Fisher in 1936 (37). The functions are used in statistical testing and the identification of two or more populations in multivariate statistical analysis, and the method reduces a complicated multivariate analysis to a

TABLE 16

ANALYSIS OF VARIANCE (AOV)

Source of Variation Due To	Sum of Squares (SS)	Degrees of Freedom	Mean Square	Calculated F-Ratio
Regression	$A'X'Y$	p	$MS(R) = \frac{A'X'Y}{p}$	
Residual	$E'E = Y'Y - A'X'Y$	n-p	$MS(E) = \frac{E'E}{n-p}$	$\frac{MS(R)}{MS(E)}$
Total	$Y'Y$	n	---	

single variate by use of linear compounds of the several variables. Rao (37) has made a comprehensive presentation of discriminant theory, and he has elaborated on its applications and limitations.

Assume that $x_1, x_2, \ldots, x_p$ are measurements which have multivariate normal distribution. In addition, N, sample size, can be classified (not in a statistical sense) into two groups of N_1 and N_2 with means of $\bar{x}_i*$ and $\bar{x}_i**$, respectively. A linear expression

$$a_1x_1 + a_2x_2 + \ldots + a_px_p \qquad (4.10)$$

where x_i are assigned mathematical variables, and a_i are coefficients (parameters) may be selected to maximize discrimination between the two groups. The variance of the linear expression above is given as (37,38):

$$V = \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j v_{ij} \qquad (4.11)$$

where v_{ij}'s are elements of the dispersion matrix.

If the differences in means are denoted by $d_i = \bar{x}_i* - \bar{x}_i**$, $(i = 1,2,\ldots,p)$, then the square of the difference in the mean values of expression (4.10) can be expressed as:

$$T = a_1d_1 + a_2d_2 + \ldots + a_pd_p \qquad (4.12)$$

$$T^2 = (a_1d_1 + a_2d_2 + \ldots + a_pd_p)^2 \qquad (4.13)$$

or

$$T^2 = \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j d_i d_j$$

The coefficients of equation (4.13) are unknown, and in order to obtain the coefficients while maximizing T^2, the following function is formed by introducing the Lagrangian multiplier (37):

$$W = T^2 - \lambda V \tag{4.14}$$

or $$W = \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j d_i d_j - \lambda \sum_{i=1}^{p} \sum_{j=1}^{p} a_i a_j v_{ij} \tag{4.15}$$

Now, differentiating W with respect to a, and after rearranging and simplifying, the following system of equations result that may be solved for a_i. Since the v_{ij}'s are the elements of the dispersion matrix which is positive definite, matrix (V'V) is a nonsingular matrix and therefore, the system of equations below can be solved for a_i's.

$$\begin{array}{l} v_{11}a_1 + v_{12}a_2 + \ldots + v_{1p}a_p = d_1 \\ v_{12}a_1 + v_{22}a_2 + \ldots + v_{2p}a_p = d_2 \\ \cdot \quad \cdot \qquad \cdot \quad \cdot \qquad \cdot \; \cdot \; \cdot \qquad \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \quad \cdot \qquad \cdot \; \cdot \; \cdot \qquad \cdot \quad \cdot \qquad \cdot \\ \cdot \quad \cdot \qquad \cdot \quad \cdot \qquad \cdot \; \cdot \; \cdot \qquad \cdot \quad \cdot \qquad \cdot \\ v_{1p}a_1 + v_{2p}a_2 + \ldots + v_{pp}a_p = d_p \end{array} \tag{4.16}$$

This concept and the method of calculation outlined were first developed by Fisher (37).

In order to summarize the discriminant procedures for classification purposes, the following steps may be outlined:

1. Determine the coefficients of discriminant function $a_1, a_2, \ldots, a_p$ from equation (4.16) by solving p simultaneous equations for p unknowns.

2. Establish a discriminant function (commonly known as the discriminant score (34)) as:

$$S = a_o + a_1 x_1 + \ . \ . \ . + a_i x_i \quad (i=1,2,\ldots,p) \quad (4.17)$$

3. Given several groups with p normally distributed variables, calculate the means in each group, and the means for each variable among the groups. Then, the linear discriminant scores $S_1, S_2, \ldots, S_p$ can be computed from formula (4.17).

4. The score obtained by the procedure above for each individual is compared with the general mean score (L), and the individual is assigned to that group for which S is maximum. The method of assigning is based on the constant likelihood ratio concept and the general mean refers to the overall mean of the discriminant score (37,38).

The test of the hypothesis of discriminant functions is quite complex, and it requires the examination of U-statistics, Mahalanobis's D^2, and F-ratio (generalized) tests. The presentation of these complicated statistical concepts is beyond the scope of this study, and elaborate derivations and discussions are given in advanced mathematical statistics texts (37,39,40); however, the application of the method and test of hypothesis will be given in later chapters.

CHAPTER V

RESEARCH METHODOLOGY

Introduction

The scientific investigation can be thought of as a product of the body of available knowledge and the body of procedures. Access to the body of knowledge permits the control of environment and the procedures are utilized to increase this body of information and knowledge. The source of data and mathematical concepts discussed in the preceeding chapters provide these two ingredients which will permit the continuation of this study.

The research methodology refers to scientific methods which will be used in the analysis of data. The logics of procedures and analyses will be presented, and the criteria for testing and selecting the proper models will be outlined.

Definitions and Assumptions

Length of Bridge--Any structure that had over twenty feet length was considered a bridge structure (30).

The length of bridge was measured from end to end as was specified on the design plans.

Defined Limits--A limit within which all information regarding bridge structure and geometrics of roadway was obtained. In order to obtain valid geometric data, five hundred feet on either side of the bridge was determined to be adequate for this purpose. The geometrics of the roadway were recorded within these defined limits except for sight distance.

Selection of Variables

The selection of response and independent variables is an important step in the planning stage of a study. During the planning stage of this study, numerous aspects of the accident as a response variable were reviewed and the feasibility of available data was assessed.

The choice of independent variables needs careful study in order to meet certain assumptions which are essential in the model building process. These assumptions are independency among variables and the normality assumption, which were indicated earlier. Having these in mind, a total of eleven independent variables were selected to be included in the analysis of this study. The symbols used and description of these variables are shown in Table 17.

The majority of independent variables under study were considered to be contributing factors to accident occurrence at bridge approaches. Previous studies (see Chapter II) indicated that the deficiency in these factors bears direct relationship to the accidents. Although the limited number of variables do not appear to be roadway elements, they were included in the analysis because their indirect effects on accident occurrence were envisioned. These independent variables are simply the potential elements, and they could be dropped out because of their insignificant contribution to the model under study.

Classification of Data

The accident data were classified into several categories for the purpose of studying various conditions on accident occurrence. This classification permitted separation of the conditions in which accidents occurred, and it also provided homogeneous data for the analysis. The entire data were classified as two-lane or four-lane rural highways. Two-lane roadways consisted of all state and US highways which were operating as two-way traffic facilities. The four-lane roadways included the Interstate Highway System

TABLE 17

DESCRIPTION OF VARIABLES

Symbol	Original Independent Variables
X_1	Average Daily Traffic Volume (vehicles/day)
X_2	Width of Bridge (feet)
X_3	Width of Approaching Pavement (feet)
X_4	Critical Stopping Sight Distance (feet)
X_5	Critical Approaching Grade (percent)
X_6	Horizontal Curvature (degree)
X_7	Height of Bridge Guard Rail (feet)
X_8	Length of Bridge (feet)
X_9	Traveling Speed (mph)
X_{10}	Number of Driveways and Intersections (number of openings)
X_{11}	Index of Sufficiency Rating

and all other four-lane facilities which had characteristics comparable to Interstate Highways.

The effect of weather conditions on accidents was investigated. The data in each roadway classification were further classified according to wet and dry pavement conditions. The significance of this classification and effects of independent variables as contributing factors were examined by the discriminant analysis. After this classification process, the linear models were applied to both data classifications.

It is well known that darkness reduces the driver's visibility and reaction time. In order to gain insight to the effect of roadway elements on accidents at nighttime, the data for each roadway classification were separated according to night and day accidents. The examination of data indicated that the number of nighttime accidents was much smaller than of daytime accidents. The discriminant analysis was also applied to this classification.

Further classifications were made according to ADT and width of bridge on two-lane highways. The linear regression analysis was applied to these categories in order to investigate the effect of bridge width and volume on accidents at bridge approaches. In addition, the data were stratified as to involvement of accidents according to single or multiple vehicle accidents. Discrimination analysis was applied to these classifications in order to study the nature of roadway elements as contributing factors to traffic accidents at bridge approaches. As a result, a total of eighteen categories were delineated in determining the effect of different aspects of roadway elements on traffic accidents at bridge locations.

Stepwise Discriminant Analysis

The application of discriminant analysis had two functions in this study. First, the method was used to discriminate between two groups for examining whether they belonged to two different population groups. Second, it separated those independent variables, in decending order of importance, which were contributing significantly to accident occurrence.

The particular approach that was used in accomplishing this analysis is called stepwise discriminant analysis. The Fortran computer program was developed by Paul Sampson (41) at the University of California, Los Angeles. The data for two groups were read into the computer and the following statistical information was calculated for each variable within a group:

1. Mean of each variable within a group
2. General mean, i.e., the overall mean combining groups one and two
3. Standard deviation for each variable within a group
4. Correlation matrix within the groups.

As the first step, the F-ratios were computed for each variable in the analysis. At each step of execution one variable was entered into the analysis of discrimination, and U-statistics and approximate F-ratio were computed for each step. In order to enter a variable into the discriminating function, the first of the three basic properties outlined below was used. This means that the first property was examined and if it statisfied the variable in question, it was entered into the function (41).

Property 1--The computed F-values were compared and the variable with the highest F-ratio was entered in the function.

Property 2--In each step the variables were divided into two disjoint sets: (a) the variables which were included in the function, and (b) the variables which were not included in the discriminating function. The entering variable was the one which, if partialed on the variables already in the discriminating function, had the highest multiple correlation with the groups.

Property 3--The ratio of variance within a group to total generalized variance was formed. The entering variable was the one that resulted in greatest decrease in the ratio above.

Furthermore, at each step the F-ratio of each variable was compared with the specified value of F at a certain level of α (for example 0.01 level). If the computed F-value for any variable at the terminating step were lower than that specified, the variable was excluded from the function. Thus, using the method above, the variables were determined, and the coefficient and constants of the discriminating function were computed. For the purpose of classification, the following statistics were calculated:

1. Discriminant function was evaluated for each case (observation).

2. Posterior probability for each case--that is, the probability of case i in group number one having come from group number two.

3. Mahalanobis' D^2-- a statistic in multivariate analysis which tests the differences in means between two or more populations where p characters of variables are involved.

The D^2 was computed for each case and it was compared with the generalized D^2 to classify each observation to the proper population class. The D^2 is based on the concept of distance, and it is used as a variance ratio to test the differences in means (37).

4. U-Statistics--The purpose of the U-statistics is to test (overall concept of classification) whether the assigned contrast agrees significantly with that from the data. This statistic was also used as a variance ratio to test the overall result of the classification.

The discriminant analysis as described above is an applicable concept such that a classification or its significant contribution to a cause, based on certain determined variables, cannot be quantified. This concept was used in the analysis of wet versus dry, night versus daytime, and multiple versus single car accidents in classifying the data for this study.

Stepwise Linear Regression Analysis

The multiple regression method used in this study is a useful mathematical tool. The model of this kind is quite suitable in studying the effect of a large number of independent variables on the response variable. As mentioned earlier, eleven potential independent variables were considered for investigation during the analysis. The large size of the sample and the inclusion of many independent variables in the analysis necessitate that the computer technique be used in building the models of this nature.

There are several different methods by which linear regression models can be built. The method used in this study is referred to as step-wise regression procedure, similar to the step-wise discriminant analysis discussed in this chapter. The analysis of stepwise regression procedure consists of adding the variables to the model step by step, based on the criterion of "goodness of fit." This procedure is based on the examination of F-ratio to determine the priority of a variable to be entered in the model.

The most outstanding characteristics of stepwise regression analysis are that the final equation contains only those variables that are significantly contributing to the reduction of residuals in the model. For example, a variable may be highly significant in entering the function at early stages, but the addition of several other variables to the equation at later stages may create a situation such that the first variable would no longer be a significant variable to be retained in the model (42). Thus, an equation is developed at each step with variables which are the most significant contributors to the improvement of the model. The intermediate equations, thus developed, have forms as:

$$
\begin{array}{lll}
Y = a_o + a_1\, x_1 & \text{Stage 1} & \\
Y = b_o + b_1\, x_1 + b_2\, x_2 & \text{Stage 2} & \\
Y = c_o + c_1\, x_1 + c_3\, x_3 & \text{Stage 3} & (5.1) \\
.\quad .\quad .\quad .\quad .\quad . & \text{Stage 4} & \\
.\quad .\quad .\quad .\quad .\quad . & &
\end{array}
$$

This procedure continues in the manner above until the F-ratio is no longer significant at a certain assumed probability error.

The computer program for processing the stepwise regression procedure was originally developed by Efroymson (42), and it was programmed in a refined manner in Fortran language by the Health Science Computing facilities of the University of California, Los Angeles (41). The data and tranformation (if any) cards were read into the computer. In order to add a variable to the model, the potential variables were examined. Then, a variable which could reduce the residual sum of squares to a minimum was selected to enter the equation. This means that the criterion for choosing a variable from among several potential variables required the entering variable to have the following properties (41).

Property 1--The entering variable is a variable that reduces the residual sum of squares to a minimum.

Property 2--It is a variable that has the highest F-ratio among the potential variables.

Property 3--The variables (including dependent variable) are divided into two disjoint sets at each step. The first set includes all independent variables in the equation, and the second set includes the dependent variable and the remaining independent variables which are not in the equation. Thus, the entering variable is a variable that has the highest partial correlation coefficient with the dependent variable when partialed in the manner described above.

The program provides extensive statistical information as listed in Table 18. This type of comprehensive program is quite useful in studying the interaction of variables in the model. In addition, the plots of residual versus the variable can be used to investigate the behavior of each variable and the feasibility of certain types of transformation.

Examination of Multiple Correlation Coefficient

There are three different types of correlation coefficients: (a) simple, (b) partial, and (c) multiple

TABLE 18

OUTPUT INFORMATION OF THE PROGRAM

Information Type	Description of Output Information
Preliminary	Input Specification (sample size, assumed probability error,etc.)
	Means for Each Variable
	Standard Deviation for Each Variable
	Covariance Matrix
	Correlation Matrix
Intermediate	Multiple Correlation Coefficient
	Standard Error of Estimate
	Analysis of Variance Table
	Regression Coefficients[a]
	Standard Error[a]
	F-ratio to Remove a Variable[a]
	Partial Correlation Coefficient[b]
	F-ratio to Enter a Variable[b]
	Tolerance Value[b]
Final	Listing of Residuals
	Plots of Residuals Versus Input Variables
	Summary Table

[a] These information are for variables in the equation.
[b] These information are for variables which are not in the equation.

correlation. The simple coefficient of correlation indicates the degree of association between two random variables x and y. The value of this coefficient (also (a) and (b)) lies between -1 and +1. If the simple coefficient of correlation is denoted by r, its value can be computed from the equation below:

$$r = \frac{\text{COV } (x,y)}{[V(x)V(y)]^{\frac{1}{2}}} \tag{5.2}$$

where: COV (x,y) = covariance of x and y

V(x) = variance of x

V(y) = variance of y

If r has a value of plus one or minus one, then x and y are correlated linearly. If x and y are independent, then the covariance will be zero indicating that x and y are uncorrelated (r = 0).

The partial correlation coefficient is used where several variables are involved in the analysis. The partial coefficient of correlation is specifically desirable when one is interested to know the correlation between one independent variable and the response variable without taking into consideration the effect of remaining independent variables in the equation (43). The coefficient of partial correlation can be calculated in terms of simple correlation coefficients as shown in equation (5.4). First, it is assumed that the regression equation is given as:

$$y = a_o + a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \tag{5.3}$$

$$r_{13.2} = \frac{r_{13} - r_{12} r_{23}}{\left[(1-r_{12}^2)(1-r_{12}^2)\right]^{\frac{1}{2}}} \tag{5.4}$$

where: $r_{13.2}$ = partial correlation coefficient between y and x_3 when x_2 is held constant.

r_{12} = simple correlation between y and x_1

r_{13} = simple correlation between y and x_3

r_{23} = simple correlation between x_2 and x_3

Referring to equation (4.7), the coefficient of multiple determination (36) is defined as the ratio of the sum of squares due to regression to the sum of squares about the mean. Denoting this ratio as R^2, the following relationship results:

$$R^2 = \frac{\text{SS due to regression}}{\text{SS about mean}} = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} \tag{5.5}$$

or, it can be expressed as a percentage which is called the "percentage of variation". Now, the coefficient of multiple correlation can be defined as the square root of R^2.

$$R = \pm \left[\frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}\right]^{\frac{1}{2}} \tag{5.6}$$

The coefficient of multiple correlation is a means of expressing the association that exists between two or more variables. The range of the correlation coefficient varies between +1 and -1, and the reason for a minus sign can be detected in equations (5.5) and (5.6). If the numerator of equation (5.6) is zero, the total variation is unexplained and the coefficient of correlation is zero. If the unexplained sum of the squares is zero, the total variation is explained and the coefficient of correlation is one. In general, the higher multiple correlation coefficient indicates the degree of adequacy of the model, but caution should be exercised in application and interpretation of the coefficient.

Special attention is required in the use of the coefficient of multiple determination and the coefficient of multiple correlation. The term R^2 is used to measure the degree of usefulness of variables in the model, and as a criterion indicates, the percentage of total variation about the mean ($\bar{y}$) that is due to regression (36). The statistician (35,36) suggests that cautions be taken in using the correlation coefficients in the decision making process. Clearly, the value of R^2 can be made to approach unity by the addition of more independent variables to the model. Introducing such new variables is very likely to raise the value of R^2, but in reality the forcing of new variables may not have any meaningful purpose toward the improvement of the model under study. The selection of a new variable must be based on experience, and it must be a variable influencing the value of the response variable in a meaningful manner. The low value of the coefficient of correlation should not indicate that the independent variables do not contribute significantly in measuring the intended quantities. This may result from inaccurracy in data or the influence of other factors. Consequently, the coefficient of correlation should not be used as the sole decision-making factor, and its low values should not be discouraging.

Examination of Residuals

In the illustration presented earlier, it was shown that the difference between the observed and estimated values (using derived equation) of the response variable was called the residual. The residual is the amount of variation which cannot be explained by the derived model. Consequently, it is desirable to examine residuals in building a linear regression model. The main purpose is to investigate violation of assumptions regarding residual terms that may arise upon the application of data.

The quickest method of detecting violation is the graphical procedure, suggested by Draper and Smith (36). The response variable and independent variables are plotted against their residual terms. Draper and Smith suggest that if the plot of residuals against its respective variable indicates a horizontal band (around zero residual line), no abnormality can be suspected. This means that, based on the data utilized, assumptions do not appear to be violated. In contrast, abnormality can be suspected for the graphs which do not have horizontal bands. For example, a trapezoidal band indicates a variable variance, an arch-type band depicts non-linearity, and finally a slanted horizontal band implies the need for inclusion of a linear term in time. Other computational methods can be used to examine the assumptions concerning residual distribution. The Durbin-Watson test (43) and the "Unit Normal Deviate" test (36) are used to check the independency and normality assumptions. The graphical method is recommended where large samples with many parameters are employed in the analysis.

The program utilized in this study was capable of providing residual plots for every variable in addition to a residual graph against estimated response variable. The residual plots were examined thoroughly during the analysis, and when abnormality was observed, the proper transformations were applied to remove the deficiency.

Selection of Best Linear Model

The determination of a best model refers to choosing a model which can satisfy the assumptions made earlier to the highest degree. Since there were eleven independent variables in this study and the ultimate goal was to develop a simple model, it was essential to employ a method that could enter only the most important variables into the model. The stepwise regression procedure, described earlier in this chapter, was used as a means of delineating the most important variables for inclusion in the model. As a result, the stepwise regression procedure provided the best model which could be developed through elimination and combination of the potential variables. Generally, a study of this nature cannot be limited to the development of a single model, and it will require detailed investigation of several models under different environmental conditions and of those models which will be resulting from transformation of the original variables. Therefore, this study resulted in many alternate equations which required a decision-making criterion for selection of the best equation. The following criteria are those that can be used to make the desired decision possible.

The first criterion used in choosing the best equation is the coefficient of multiple determination (R^2) which was discussed earlier in this chapter.

The second method is to test the overall regression model by use of the F-ratio test. Draper and Smith (36) suggest that the ratio (mean square ratio) can be formed as:

$$F = \frac{SS\ (R/b_o)\ /\ (p - 1)}{SS\ (\text{Residual})\ /\ (n-p)} \qquad (5.7)$$

with (p-1) and (n-p) degrees of freedom where SS (R/b_o) = SS (Regression) after allowance made for b_o(SS(R/b_o) = SS due to Regression - SS due to b_o). Assume that the ratio in equation (5.7) is computed and compared with the value of F obtained from the F distribution tables at a certain level of risk α. If the computed value exceeds the table value, the derived regression equation is statistically valid at the specified risk level. This means that the proportion of variation of the observed data which accounted for by the derived model is greater than would be expected by chance (assuming identical condition for the data) alone at a certain risk level.

The F-ratio test is not always a good testing procedure in examining the predictive models. Draper and Smith (36) make reference to Wetz's work concerning the use of

F-ratio test. Wetz's work suggests that in order for a model to be a good predictor it should have a computed F-ratio about four times higher than that obtained from table of F-distribution at a certain level of risk. For example, if $F(10, 60, 0.01) = 2.63$ is selected from the table of F-distribution, the computed value of F-ratio should exceed $4 \times 2.63 = 10.52$. Therefore, a predictive model should have an F-value of 10.52 in order to be a satisfactory model.

The standard error of estimate is another means to examine the predictive power of the derived model. The standard error of estimate is the square root of residual mean square, and it is the estimate of the variance about the regression. The low values of standard error of estimate indicate that the model in question is capable of estimating the response variable more precisely than the model that has a higher standard error of estimate. Since the value of the standard error of estimate can be reduced to zero by adding more variables to the model, caution should be exercised in using this criterion for selection purposes.

In general, any of the methods described can be employed for comparison of several derived models and these methods as decision making criteria can provide reasonable results. The method used in this study for selecting the proper models was a combination of two test criteria, and it consisted of an examination of F-ratio and the coefficient of multiple determination. The F-ratio test and the coefficient of multiple determination were used to test the adequacy and percentage of variation that could be explained by the models, respectively.

CHAPTER VI

INTERPRETATION OF ANALYSES

Introduction

The analysis of the study was organized in two different stages, namely preliminary (experimental), and final. The preliminary stage of the research consisted of planning the model and running experimental programs to investigate the effect of the independent variables on the response variable. The final stage of the analysis included the implementation of the experience gained from stage one (such as transformations and omission of variables) and production run of the classified models at a specified level of F-ratio.

The implications of the analysis and the critical review of the results obtained are essential in any type of research work. Thus, the models developed through this study will be examined in detail regarding their accuracy and usage. In addition, the applications of the resulting models will be illustrated and the limitation of data used in this study will be discussed.

Characteristics of Data

The data were classified into two classifications, as two-lane and four-lane highways. The means and standard deviations for each variable were computed as shown in Tables 19 and 20 for two-lane and four-lane highways, respectively. The examination of standard deviations indicated a large variation in the data. This variation may be the result of different levels of standards which have been utilized in construction of roadways through past decades. For example, the grades or curvatures on the roadways which were built between 1940 and 1950, generally, were much higher than those built in recent years.

A comparison of computed data was made between two-lane and four-lane roadways (Tables 19 and 20). This review indicated that the mean accident rates of four-lane highways on bridge approaches are higher than for two-lane roadways. The mean accident rates of two-lane roadways was about one-half of accident rates of four-lane roadways for the calendar years 1968 and 1969. The comparison of mean stopping sight distances showed that the four-lane roadways have shorter stopping sight distances than two-lane highways. Although the difference in the two sight distances is not significant, the reason for shorter sight distances on four-lane roadways

TABLE 19

MEAN AND STANDARD DEVIATION OF VARIABLES FOR OVERALL DATA[a]
TWO-LANE ROADWAYS

Variable		Mean	Standard Deviation
Y:	Accident Per Year	1.632	1.904
X(1)[b]	: ADT	2494.021	2225.476
X(2)	: Width of Bridge	26.504	8.342
X(3)	: Width of Pavement	22.536	2.273
X(4)	: Stopping Sight Distance	1262.193	709.139
X(5)	: Approaching Grade	-0.104	3.356
X(6)	: Degree of Curvature	1.478	2.923
X(7)	: Height of Guard Rail	3.435	0.987
X(8)	: Length of Bridge	293.530	569.837[c]
X(9)	: Speed Before Contact	52.866	12.012
X(10)	: Number of Openings	1.604	2.426
X(11)	: Index of Sufficiency Rating	80.754	35.223

[a] Based on a sample size of 611 distinct locations.

[b] The numbers in parentheses indicate subscripts of X's.

[c] The standard deviations are higher than the means because of a large variation in the data and the negative range has no physical meaning.

TABLE 20

MEAN AND STANDARD DEVIATION OF VARIABLES FOR OVERALL DATA[a]
FOUR-LANE ROADWAYS

Variable		Mean	Standard Deviation
Y:	Accident Per Year	3.221	3.455
X(1)[b]	: ADT	15773.895	12560.621
X(2)	: Width of Bridge	33.691	9.486
X(3)	: Width of Pavement	24.103	3.867
X(4)	: Stopping Sight Distance	1129.228	864.992
X(5)	: Approaching Grade	-0.131	2.450
X(6)	: Degree of Curvature	0.310	0.901
X(7)	: Height of Guard Rail	3.533	0.756
X(8)	: Length of Bridge	315.294	554.570[c]
X(9)	: Speed Before Contact	56.485	11.854
X(10)	: Number of Openings	0.993	1.453
X(11)	: Index of Sufficiency Rating	91.118	10.613

[a] Based on a sample size of 136 distinct locations.

[b] The numbers in parentheses indicate subscripts of X's.

[c] The standard deviations are higher than the means because of a large variation in the data and the negative range has no physical meaning.

is not known.

The correlation matrices were also computed for each classification of the roadways mentioned as shown in Tables 21 and 22. Examination of the correlation matrix for two-lane roadways indicated that a relatively high correlation exists between the Index of Sufficiency Rating and several other independent variables: width of bridge, width of approaching pavement, and the number of openings. The high relative correlation may be attributed to the fact that the independent variables mentioned were rated and included in calculating the Sufficiency Rating Index (see Chapter III). The intercorrelation of the Index of Sufficiency Rating and the width of bridge is quite high (approximately 0.70) for four-lane highways, and the correlation is moderate between the Index of Sufficiency Rating and the width of the approaching pavement. The intercorrelation of independent variables in the analysis of regression models is not a desirable characteristic. This type of intercorrelation will not allow the effect of other independent variables (i.e., width of bridge, width of approaching pavement) to be reflected in the model.

TABLE 21

CORRELATION MATRIX
TWO-LANE ROADWAYS

Vari-ables	X(01)[a]	X(02)	X(03)	X(04)	X(05)	X(06)	X(07)	X(08)	X(09)	X(10)	X(11)	X(12)
X(01)	1.000	0.212	-0.102	-0.082	-0.036	0.082	0.132	0.052	0.041	-0.075	-0.066	-0.101
X(02)		1.000	0.350	0.231	0.064	0.021	-0.009	0.037	0.044	-0.109	0.078	0.197
X(03)			1.000	0.368	0.105	0.052	-0.195	-0.211	-0.052	0.025	0.234	0.426
X(04)				1.000	0.113	0.042	-0.221	-0.042	0.144	0.018	0.340	0.438
X(05)					1.000	0.039	-0.051	0.003	-0.046	-0.022	0.032	0.015
X(06)						1.000	-0.022	0.042	-0.040	0.014	-0.090	0.042
X(07)							1.000	0.064	-0.041	-0.105	-0.039	-0.134
X(08)								1.000	0.164	-0.024	-0.138	-0.063
X(09)									1.000	-0.073	0.007	0.154
X(10)										1.000	-0.056	0.024
X(11)											1.000	0.352
X(12)												1.000

[a] The numbers in parentheses indicate subscripts of X's.

TABLE 22

CORRELATION MATRIX
FOUR-LANE ROADWAYS

Variables	X(01)[a]	X(02)	X(03)	X(04)	X(05)	X(06)	X(07)	X(08)	X(09)	X(10)	X(11)	X(12)
X(01)	1.000	0.615	0.102	0.174	-0.199	0.067	-0.048	0.236	0.096	-0.276	-0.085	0.036
X(02)		1.000	0.158	0.102	-0.309	-0.027	-0.129	0.431	-0.077	-0.339	-0.119	0.122
X(03)			1.000	0.361	-0.090	-0.012	-0.010	0.198	-0.143	-0.212	0.130	0.704
X(04)				1.000	-0.079	0.000	-0.111	0.109	-0.024	-0.081	-0.129	0.400
X(05)					1.000	0.040	-0.028	-0.203	0.161	0.151	-0.037	0.057
X(06)						1.000	0.091	-0.069	-0.024	-0.042	0.046	-0.046
X(07)							1.000	-0.085	0.019	-0.008	0.058	-0.108
X(08)								1.000	-0.144	-0.162	0.067	0.078
X(09)									1.000	-0.001	-0.111	-0.136
X(10)										1.000	0.107	-0.025
X(11)											1.000	-0.027
X(12)												1.000

[a] The numbers in parentheses indicate subscripts of X's.

General Linear Regression Models

The models in this class were developed through the use of the total data without any classifications. The first group of general models was based on using eleven independent variables, and the second group of equations was derived by employing nine independent variables in the analysis. The general format of the models investigated is shown in Table 23. The models obtained for each classification of the roadway are discussed in the following sections.

Bridge Approaches on Two-Lane Highways

The analysis of accidents at two-way bridge approaches consisted of the stepwise linear regression treatment of data. The sample size for this class was 611 distinct locations on two-lane roadway system throughout the State of Oklahoma.

During stage one, several models were developed by treating the number of accidents as a response variable with the independent variables as described in the previous chapters. The models which were developed in this stage are shown in Table 24. The coefficients of multiple determination calculated for each model was not satisfactory (too low) for prediction purposes. Consequently, the examination

TABLE 23

GENERAL FORMAT OF THE MODELS INVESTIGATED

Model Number	Model Type
(01)	$Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n$
(02)	$Y = a_0 + a_1X_1 + \ldots + a_nX_n$
(03)	$\frac{1}{Y} = a_0 + a_1X_1 + \ldots + a_nX_n$
(04)	$\ln Y = a_0 + a_1X_1 + \ldots + a_nX_n$
(05)	$\ln Y = a_0 + a_1X_1 + \ldots + a_nX_n + b_1\ln X_1 + b_2\ln X_2 + \ldots + b_n\ln X_n$
(06)	$\ln Y = a_0 + a_1\ln X_1 + \ldots + a_n\ln X_n$
(07)	$Y = a_0 + a_1X_1 + (a_2X_1X_2 + \ldots + a_nX_1X_n)$
(08)	$Y^a = a_0 + a_1X_1 + (a_2X_1X_2 + \ldots + a_nX_1X_n)$
(09)	$(Y)^{\frac{1}{2}} = a_0 + a_1X_1 + (a_2X_1X_2 + \ldots + a_nX_1X_n)$
(10)	$Y = a_0 + a_1X_1 + \ldots + a_nX_n + b_1X_1^2 + b_2X_1^2X_2 + b_3X_1^2X_3 + b_4X_2X_3 + b_5X_3X_4 + b_6X_4X_5 + b_7X_5X_6$
(11)	$\ln Y = a_0 + a_1X_1 + \ldots + a_nX_n + b_1X_1^2 + b_2X_1^2X_2 + b_3X_1^2X_3 + b_4X_2X_3 + b_5X_3X_4 + b_6X_4X_5 + b_7X_5X_6$

TABLE 23--Continued

Model Number	Model Type
(12)	$Y/X_1 = a_0 + a_1X_1 + . . . + a_nX_n + b_1X_1^2 + b_2X_1^2X_2 + b_3X_1^2X_3 + b_4X_2X_3 + b_5X_3X_4 + b_6X_4X_5 + b_7X_5X_6$
(13)	$\ln(Y/X_1) = a_0 + a_1X_1 + . . . + a_nX_n + b_1X_1^2 + b_2X_1^2X_2 + b_3X_1^2X_3 + b_4X_2X_3 + b_5X_3X_4 + b_6X_4X_5 + b_7X_5X_6$
(14)	$\log_{10}(Y/X_1) = a_0 + a_1X_1 + . . . + a_nX_n + b_1X_1^2 + b_2X_1^2X_2 + b_3X_1^2X_3 + b_4X_2X_3 + b_5X_3X_4 + b_6X_4X_5 + b_7X_5X_6$
(15)	$\ln(Y/X_1) = a_0 + a_1X_1 + . . . + a_nX_n + X_1^2 (b_0 + b_1X_2 + . . . + b_nX_n + 1)$
(16)	$\ln(Y/X_1) = a_0 + a_1X_1 + X_1^2(b_0 + b_1X_2 + . . . + b_nX_n + 1)$
(17)	$\log_{10}(Y/X_1) = a_0 + a_1X_1 + . . . + a_nX_n + b_1X_1^2 + b_2X_2X_3 + b_3X_3X_5 + b_4X_3X_6 + b_5X_2X_8 + b_6X_5X_6 + b_7X_4X_6 + b_8X_4X_7 + b_9X_2X_5 + b_{10}X_2X_6 + b_{11} (X_3- X_2)$

TABLE 23 --Continued

Model Number	Model Type
(18)	$\log_{10}(Y/X_1) = a_0 + a_1X_1 + a_2(1/X_2) + a_3(1/X_3) + b_4(1/X_4) + a_5X_5 + a_6X_6 + a_7X_7 + a_8X_8 + a_9X_9 + a_{10}X_1^2$
(19)	$\log_{10}(Y/X_1) = a_0 + a_1X_1 + \ldots + a_nX_n + b_0X_1^2 + b_1(X_3/X_2) + b_2X_1X_5 + b_3X_1(X_3/X_2) + b_4X_2^2 + b_5X_7^2$
(20)	$(Y)^{1/2} = a_0 + a_1(X_1)^{1/2} + \ldots + a_n(X_n)^{1/2}$

[a]A linear regression analysis with ten variables only.

of residual graphs was necessary to investigate the cause of abnormality and variation in the data. It was found that non-linearity and variable variance existed in the collected data. In order to remove the effect of non-linearity from the data, several transformations were performed on response and independent variables. Although some of these transformations improved the accuracy of the model to some extent, there was no significant success toward the improvement of the predictability of the model.

Further investigation of the variables and correlation matrix indicated that a relatively high inter-correlation existed between Index of Sufficiency Rating and three independent variables. In addition, it was found that the distribution of the "Speed Before Contact" did not conform to characteristics of a normal distribution. Therefore, it was decided that variables number ten (Speed Before Contact) and eleven (Index of Sufficiency Rating) be excluded from the analysis. The linear regression was performed on the remaining variables and showed no substantial improvement in the models. These models are shown in Table 24.

The study of the dependent variable (number of accidents per year) indicated that the number of accidents conforms closely to Poisson distribution (41,42). Since the

TABLE 24

MODELS RESULTING FROM OVERALL DATA

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(01)	$Y = 0 + 0.00023X_1 - 0.02927X_2 + 0.09504X_3 + 0.04848X_4 + 0.08420X_6 - 0.00554X_{11}$	0.6865	0.4713	1.8301	90.038
(02)	$Y = 2.22517 + 0.00024X_1 - 0.03457X_2 + 0.05111X_5 + 0.06214X_6 - 0.00447X_{11}$	0.3211	0.1031	1.8109	13.914
(03)	$1/Y = 0.66187 - 0.00003X_1 - 0.00969X_6 - 0.00003X_8 + 0.00310X_{11}$	0.2890	0.0835	0.2700	13.811
(04)	$\ln Y = 0.70867 + 0.00006X_1 + 0.01004X_5 + 0.01736X_6 + 0.00007X_8 - 0.00754X_{11}$	0.3509	0.1232	0.4783	16.995
(05)	$\ln Y = 0.09197 + 0.18312\ln X_1 + 0.00985\ln X_5 + 0.01878\ln X_6 + 0.04293\ln X_8 + 0.12056\ln X_9 - 0.43150\ln X_{11}$	0.3762	0.1415	0.4736	16.592

TABLE 24--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(06)	$Y = 0.78816 + 0.00218X_1 + 0.00000X_1X_2 - 0.00000X_1X_4 + 0.00002X_1X_5 + 0.00001X_1X_6 - 0.00003X_1X_7 - 0.00002X_1X_{11}$	0.5534	0.3063	1.5953	38.036
(07)	$Y = 0.81082 + 0.00142X_1 - 0.00001X_1X_2 - 0.00003X_1X_3 - 0.00000X_1X_4 + 0.00004X_1X_5 + 0.00003X_1X_6 + 0.00001X_1X_9$	0.4311	0.1859	1.7282	19.666
(08)	$(Y)^{\frac{1}{2}} = 1.01631 + 0.32084X_1 - 0.00241X_1X_3 - 0.00779X_1X_3 - 0.00001X_1X_4 + 0.00575X_1X_5 + 0.00711X_1X_6 + 0.00384X_1X_9$	0.4240	0.1798	0.3822	18.879
(09)	$Y = 22.83974 + 0.09448X_1 - 1.00982X_2 - 0.92061X_3 + 0.06129X_6 + 0.36332X_1^2 - 0.00117X_1^2X_3 + 0.04269X_2X_3 + 0.01386X_5X_6$	0.4624	0.2138	1.7011	18.160

TABLE 24--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(10)	$\ln Y = 0.25368 + 0.12668X_1 + 0.02619X_6 - 0.00019X_1^2X_3 - 0.00043X_2X_3 + 0.00317X_5X_6$	0.3888	0.1512	0.4706	21.550
(11)	$Y/X_1 = 0.00156 - 0.00051X_1 + 0.00002X_5 + 0.00005X_6 + 0.00006X_1^2 - 0.00000X_1^2X_2 + 0.00000X_2X_3$	0.5284	0.2792	0.0009	38.998
(12)	$\ln(Y/X_1) = -6.36372 - 0.44981X_1 + 0.02708X_6 + 0.03497X_1^2 - 0.00046X_1^2X_2$	0.7405	0.5484	0.5480	183.960
(13)	$\log_{10}(Y/X_1) = -2.76381 - 0.19527X_1 + 0.01177X_6 + 0.01516X_1^2 - 0.00020X_1^2X_2$	0.7405	0.5483	0.2380	183.892
(14)	$\log_{10}(Y/X_1) = -2.96466 - 0.15855X_1 + 0.00866X_6 + 0.00434X_1^2 + 3.37783(\frac{1}{X_2}) + 22.02321\frac{1}{X_3}$	0.7336	0.5382	0.2409	140.998

TABLE 24--<u>Continued</u>

Number	Equation	R	R^2	SEE	F-Ratio
(15)	$\log_{10}(Y/X_1) = -2.94314 - 0.15986X_1 + 0.01085X_6 + 0.00435X_1^2 + 0.15403\frac{X_4}{X_3}$	0.7311	0.5345	0.2416	173.970
(16)	$\log_{10}(Y/X_1) = -2.80456 - 0.16417X_1 + 0.05034X_6 + 0.00003X_8 + 0.00468X_1^2 + 0.00143X_5X_6 - 0.00171(X_4-X_3)$	0.7363	0.5422	0.2400	119.224
(17)	$(Y)^{\frac{1}{2}} = 0.19779 + 0.11583(X_1)^{\frac{1}{2}} - 0.05432(X_2)^{\frac{1}{2}} + 0.00414(X_5)^{\frac{1}{2}} + 0.02984(X_6)^{\frac{1}{2}} + 0.00141(X_8)^{\frac{1}{2}}$	0.3668	0.1346	0.2064	18.816

[a] Standard Error of Estimate

linear regression analysis assumes normality as a basic assumption in the derivation of the models, the response variable was converted to rate as accident per thousand of vehicles. Furthermore, the low statistics obtained in the initial stages of analyses indicated the necessity of including other variables in the models in addition to the principal variables. Thus, using the technique suggested in reference (36) variables such as cross-products, squares, and other combination of independent variables were added to the models. These modifications, combined with logarithmic transformation of the response variable, resulted in a good fit of the data to the proposed models. These techniques are reflected in model numbers 11, 12, and 13 as shown in Table 24. The transformation of variables removed nonlinearity, and cross-products of independent variables accounted for the way the independent variables interacted on the response variable (36).

The study of these models shows that volume of traffic (X_1), width of bridge (X_2), and degree of curvature (X_6) at bridge approaches are primary factors contributing to accident events. The approaching grade is included in model number 11 in addition to other factors mentioned above. The coefficients of multiple determination for the three models mentioned above are 27.92, 54.85, and 54.83 percent, respectively. Model number 13 has a coefficient of correlation equal to 74.05 percent with the minimum standard error of estimate 0.2380 as compared to model numbers 11 and 12. The variables were entered into the equation by use of sequential F-test of hypothesis at a probability error level of 0.05 or lower. Model number 13 can be considered to be the best predictor model for this class of data.

Bridge Approaches on Four-Lane Highways

The analysis of bridge approaches on four-lane roadways was quite similar to that of two-lane roadways as discussed in the previous section. The sample size used was 136 distinct locations on four-lane roadway systems throughout the State of Oklahoma.

The analysis consisted of fitting twenty different models to the data in this class. Although the elementary models with few transformations produced a good fit, attention was directed toward more complicated models in order to obtain a higher coefficient of multiple determination. The results of stepwise regression analysis indicated that the applications of complicated models did not improve the accuracy of the estimates and the coefficient of multiple determination decreased considerably. The resultant models for bridge approaches on four-lane highways are shown in Table 25. The study of these models showed that model number 21 is the most reliable equation representing the data in this class. The coefficient of multiple correlation is 71.52 percent with a standard error of estimate of 0.5900. The variables in the model explain approximately 51.16 percent of the variation of accidents per year.

The most significant variables for bridge approaches on four-lane highways are: the average daily traffic volume, length of bridge, width of approaching pavement, degree of curvature, and speed before contact. The testing of hypothesis was based on sequential F-tests at the approximate level of 0.05 or lower.

A comparison of model number 13 for two-lane and model number 21 for four-lane roadways indicates that the contribution of roadway elements to traffic accidents at bridge approaches is different for two-lane and four-lane facilities. For example, the accident rate on two-way bridges is dependent on volume of traffic, degree of curvature, and width of bridge. In contrast, the accident rate on one-way bridges is a function of traffic volume, width of approaching pavement, percentage of grade, length of

TABLE 25

MODELS RESULTING FROM OVERALL DATA

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(18)	$Y = 0 + 0.00018X_1 + 0.00098X_8$	0.8237	0.6785	2.6928	141.385
(19)	$Y = -\ 2.19619 + 0.00017X_1 + 0.10231X_3 + 0.00091X_8$	0.6414	0.4114	2.6806	30.750
(20)	$\frac{1}{Y} = 0.80467 - 0.00002X_1 - 0.01009X_3 + 0.00005X_4 - 0.02098X_5 - 0.00014X_8 + 0.00456X_9$	0.6860	0.4706	0.2711	19.115
(21)	$\ln Y = -\ 0.11629 + 0.00004X_1 + 0.02669X_3 + 0.03992X_5 + 0.00028X_8 - 0.00846X_9$	0.7152	0.5116	0.5900	27.231
(22)	$\ln Y = -\ 0.95660 + 0.59542\ln X_1 + 0.59126\ln X_2 + 0.87859\ln X_3 + 0.04098\ln X_5 + 0.15679\ln X_8 - 2.10522\ln X_{11}$	0.6730	0.4529	0.6268	17.796
(23)	$Y = 0.49030 + 0.00001X_1X_3 + 0.00000X_1X_8$	0.6459	0.4172	2.6571	47.614

TABLE 25--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(24)	$Y = 0.49034 + 0.00601X_1X_3 + 0.00000X_1X_8$	0.6459	0.4172	2.6571	47.614
(25)	$(Y)^{\frac{1}{2}} = 0.96850 + 0.00146X_1X_3 + 0.00002X_1X_8$	0.6820	0.4651	0.5716	57.830
(26)	$Y = 0.36092 + 0.00091X_8 + 0.00016X_1^2X_3 + 0.00120X_2X_3$	0.6407	0.4105	2.6825	30.645
(27)	$\ln Y = -0.29934 + 0.04361X_1 + 0.00031X_8 + 0.00036X_2X_3 + 0.00004X_4X_5$	0.7071	0.5000	0.5946	32.745
(28)	$Y/X_1 = 0.00050 - 0.00003X_1 - 0.00000X_4 + 0.00001X_5 + 0.00003X_6 + 0.00000X_8 + 0.00000X_1^2X_3$	0.5874	0.3451	0.0002	11.328
(29)	$\ln(Y/X_1) = -7.69477 - 0.10770X_1 - 0.00011X_4 + 0.05545X_5 + 0.09658X_6 + 0.00029X_8 + 0.00009X_1^2X_3$	0.5746	0.3301	0.6190	10.595

TABLE 25--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(30)	$\log_{10}(Y/X_1) = -3.34177 - 0.04678X_1 - 0.00005X_4 + 0.02408X_5 + 0.04194X_6 + 0.00012X_8 + 0.00004X_1^2X_3$	0.5746	0.3301	0.2688	10.595
(31)	$\log_{10}(Y/X_1) = -5.14965 - 0.03795X_1 + 0.03779X_3 - 0.00013X_4 + 0.02173X_5 + 0.04370X_6 + 0.00013X_8 + 0.00072X_1^2 + 26.57864(1/X_3) - 153.78497(\frac{1}{X_4})$	0.6306	0.3977	0.2579	9.245
(32)	$\log_{10}(Y/X_1) = -3.32054 - 0.04558X_1 - 0.00006X_4 + 0.02626X_5 + 0.00012X_8 + 0.00086X_1^2$	0.5511	0.3037	0.2730	11.338
(33)	$\log_{10}(Y/X_1) = -3.56233 - 0.04556X_1 + 0.04828X_6 + 0.000548X_8 + 0.00089X_1^2 + 0.00019X_2X_3 + 0.00106X_3X_5 - 0.00001X_2X_8$	0.6025	0.3630	0.2632	10.420

TABLE 25--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(34)	$(Y)^{\frac{1}{2}} = -3.4227 + 0.15413(X_1)^{\frac{1}{2}} + 0.01805(X_5)^{\frac{1}{2}} + 0.00769\ (X_8)^{\frac{1}{2}}$	0.6848	0.4689	0.2651	38.846

[a] Standard Error of Estimate

bridge, and the existence of interchange in the vicinity. Although these two models are not developed identically, this trend, generally, has been observed in comparing other models developed in this study.

Classified Regression Models

The classified models in this study refer to the models which were developed from grouping the data into several classes in order to study the effect of specific environmental conditions on traffic accidents at bridge approaches. The data were grouped according to day, night, dry pavement, wet pavement, width of bridge, average daily traffic, single car, multiple car, and type of pavement surface (bituminous or concrete). The classification was also considered as passenger car or truck traffic accidents, but insufficient data in the truck group made this type of stratification infeasible.

The discriminant analysis was performed on the grouped data prior to the application of linear regression analysis. The groups considered for the analysis were: day versus night, dry versus wet pavement, and single versus multiple car traffic accident. The best discriminating functions and pertinent statistics for each case are shown in Appendix A. The coefficients of discriminant functions were significant at 0.01 level of probability error. In addition, it was found that the assigned classification was satisfactory, and U-statistics did not substantiate the classification obtained through discriminant analysis. The discriminant analysis failed when it was applied to four-lane roadways. The investigation showed that the calculated F-values did not exceed the specified values of F as 2.45 for inclusion of variables. Consequently, the linear regression analyses were performed on all groups and the results are discussed below.

Day and Nighttime Traffic Accidents

The analysis of linear regression was based on a sample size of 365 daytime and 340 nighttime traffic accidents on two-lane roadways. The experience gained in the first portion of the analysis regarding the nature of the data did not warrant the application of the data to all of the proposed models. As a result, the investigation consisted of examination of several models which were expected to produce a good fit. The equations derived are shown in Tables 26 and 27. Examination of the first four equations clearly indicates that non-linearity existed in the data, and the removal of

TABLE 26

DAYTIME TRAFFIC ACCIDENTS

TWO-LANE ROADWAY

Number	Equation	R	R^2	SEE[a]	F-Ratio
(35)	$Y = 3.64404 + 0.00023X_1 + 0.04288X_5 - 0.03544X_{11}$	0.4031	0.1625	1.5557	23.349
(36)	$\ln Y = -0.01040 + 0.18475\ln X_1 - 0.27257\ln X_2 - 0.07402\ln X_4 + 0.02394\ln X_6 + 0.03868\ln X_8$	0.3998	0.1596	0.4315	13.659
(37)	$\ln Y = 0.00468 + 0.25163X_1 - 0.00167X_1X_2 - 0.00526X_1X_3 - 0.00002X_1X_4 + 0.01264X_1X_6$	0.4449	0.1980	0.4216	17.721
(38)	$\ln(Y/X_1) = -6.25180 - 0.63767X_1 + 0.03394X_1X_6 + 0.06788X_1^2 - 0.00026X_1^2X_2 - 0.00065X_1^2X_3$	0.8092	0.6548	0.4729	136.213

TABLE 26--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(39)	$\log_{10}(Y/X_1) = -\ 2.63826 - 0.26336X_1 + 0.01298X_6 + 0.01789X_1^2 - 0.00016X_2X_3$	0.8033	0.6453	0.2079	163.747

[a] Standard Error of Estimate

TABLE 27

NIGHT TIME TRAFFIC ACCIDENTS

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(40)	$Y = 1.24594 + 0.00008X_1 - 0.01225X_2 + 0.00012X_4$	0.2773	0.0769	0.5934	9.332
(41)	$\ln Y = 0.12068 + 0.07005\ln X_1 - 0.11894\ln X_2 + 0.01078\ln X_5 + 0.02994\ln X_8$	0.2418	0.0585	0.3271	5.202
(42)	$\ln Y = 0.07395 + 0.05803X_1 - 0.00130X_1X_2 - 0.00001X_1X_4 + 0.00379X_1X_5$	0.2510	0.0630	0.3263	5.629
(43)	$\ln(Y/X_1) = -6.04264 - 0.80586X_1 + 0.07379X_1^2 - 0.00043X_1^2X_2 + 0.00000X_3X_4$	0.8682	0.7537	0.3889	256.265
(44)	$\log_{10}(Y/X_1) = -2.59442 - 0.31732X_1 + 0.00003X_4 + 0.02154X_1^2 - 0.00011X_2X_3$	0.8605	0.7405	0.1734	238.962

[a] Standard Error of Estimate

non-linear characteristics of the data is reflected in the last two equations.

The study of the equation numbers 43 and 44 in Table 27 reveals that nighttime traffic accidents at bridge approaches (two-way bridge) have a strong relationship to roadway elements. The highest multiple correlation obtained is 86.82 percent. The variables in the equation explain 75.37 percent of the variation in traffic accidents at two-way bridge approaches.

The application of the methods described in the previous chapter for the selection of the best equation indicate that equation number 44 is the best predictor of traffic accidents at nighttime on two-way bridge approaches. The multiple correlation coefficient is 86.05 percent and the variables explain 74.05 percent of variation in traffic accident on two-way bridges at nighttime. The significant variables, in the order of importance, are: volume of traffic, cross-product of width of bridge and width of pavement, and the stopping sight distance. The significance of variables included in this equation was tested by sequential F-test using probability type I error of 0.05.

It is interesting to note that a similar equation, number 39, derived for daytime traffic accidents showed a coefficient of multiple correlation of 80.33 percent with significant variables being volume of traffic, cross-product of width of bridge and width of pavement, and degree of curvature. Comparison of these two equations indicates that sight distance available at night is an important factor while the degree of curvature becomes a critical element in traffic accidents during daytime hours.

Similar equations were developed for night and daytime accidents on four-lane roadways. The results of regression analyses are shown in Tables 28 and 29. Traffic accidents on one way bridge approaches did not show relatively high correlation with roadway elements. The highest coefficient of multiple correlation obtained was 61.93 percent and the variables explain only 38.35 percent of the variation in daytime traffic accidents. In contrast, the traffic accidents showed an excellent correlation with roadway geometrics at one-way bridge approaches at nighttime. The coefficient of multiple correlation obtained for nighttime accidents, equation number 54, was 79.95 percent. The variables explain 63.93 percent of variation in nighttime traffic accidents. The significant variables included are volume of traffic and the degree of curvature. The variables were included according to sequential F-test at 0.05 level of significance.

TABLE 28

DAYTIME TRAFFIC ACCIDENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(45)	$Y = 3.33455 + 0.00011X_1 - 0.04731X_{11}$	0.6373	0.4062	2.0706	35.913
(46)	$\ln Y = -5.48661 + 0.57396\ln X_1 + 0.13789\ln X_8$	0.6434	0.4140	0.5932	37.084
(47)	$\ln Y = -0.05160 + 0.00141X_1X_3 + 0.00002X_1X_8$	0.7097	0.5037	0.5458	53.291
(48)	$\ln(Y/X_1) = -8.04423 - 0.09564X_1 + 0.00025X_8 - 0.00007X_1^2X_3$	0.5865	0.3440	0.5730	18.178
(49)	$\log_{10}(Y/X_1) = -3.63452 - 0.04589X_1 + 0.00011X_8 + 0.00087X_1^2 + 0.00019X_2X_3 + 0.00004X_4X_6$	0.6193	0.3835	0.2436	12.689

[a] Standard Error of Estimate

TABLE 29

NIGHT TIME TRAFFIC ACCIDENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(50)	$Y = 1.02116 + 0.00006X_1$	0.3896	0.1518	1.2972	15.210
(51)	$\ln Y = -1.99969 + 0.19238\ln X_1 + 0.09795\ln X_8$	0.3618	0.1309	0.5305	6.324
(52)	$\ln Y = 0.05156 + 0.00047X_1X_3 + 0.00001X_1X_8$	0.4192	0.1757	0.5167	8.953
(53)	$\ln(Y/X_1) = -7.62505 - 0.15687X_1 + 0.14055X_6 + 0.00254X_1^2$	0.7995	0.6392	0.5419	49.015
(54)	$\log_{10}(Y/X_1) = -3.31152 - 0.06813X_1 + 0.06104X_6 + 0.00110X_1^2$	0.7995	0.6392	0.2353	49.012

[a] Standard Error of Estimate

Dry and Wet Pavement Traffic Accidents

The environmental and human factors are affected as the result of changes in weather conditions. In order to study the contribution of roadway elements to traffic accidents under wet pavement condition, the data were grouped according to dry and wet types of accidents with sample sizes of 476 and 215, respectively. The application of linear regression analysis, to the stratified data for two-way bridges, produced a number of equations which are shown in Tables 30 and 31.

The examination of equations derived for dry pavement conditions indicates that equation number 59 is the best predictor of traffic accidents at bridge approaches with two-way traffic. The coefficient of multiple correlation is 80.23 percent with standard error of estimate of 0.2063. The coefficient of multiple determination calculated for this equation is 64.37 percent and indicates that approximately 64 percent of variation in traffic accidents can be explained by the five variables in the equation.

The investigation of wet pavement conditions revealed that the best equation describing traffic accidents in this class is the model number 63. The computed R value for this equation is 86.21 percent with the standard error of

TABLE 30

DRY PAVEMENT TRAFFIC ACCIDENTS

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(55)	$Y = 1.6293 + 0.00018X_1 - 0.01795X_2 + 0.02462X_5 + 0.03933X_6 - 0.00297X_{11}$	0.3809	0.1451	0.9729	15.957
(56)	$\ln Y = - 0.31054 + 0.15907\ln X_1 - 0.21131\ln X_2 + 0.00984\ln X_5 + 0.01914\ln X_6$	0.3367	0.1134	0.4083	15.055
(57)	$\ln Y = 0.04692 + 0.18658X_1 - 0.00120X_1X_2 - 0.00500X_1X_3 + 0.00001X_1X_4 + 0.0057X_1X_5 + 0.00002X_1X_8 + 0.01030X_1X_6$	0.4054	0.1644	0.3976	13.151
(58)	$\ln(Y/X_1) = - 6.13347 - 0.67209X_1 + 0.01047X_5 + 0.02293X_6 + 0.06390X_1^2 - 0.00015X_1^2X_2 - 0.00055X_1^2X_3$	0.8078	0.6526	0.4695	146.820

TABLE 30--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(59)	$\log_{10}(Y/X_1) = -\ 2.61542 - 0.27551X_1 + 0.00471X_5 + 0.00827X_6 + 0.01811X_1^2 - 0.00011X_2X_3$	0.8023	0.6437	0.2063	169.849

[a] Standard Error of Estimate

TABLE 31

WET PAVEMENT TRAFFIC ACCIDENTS

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(60)	$Y = 2.75885 + 0.00015X_1 + 0.04407X_5 - 0.02313X_{11}$	0.3314	0.1098	1.3063	8.675
(61)	$\ln Y = 0.26227 + 0.11127\ln X_1 - 0.29619\ln X_2 - 0.21667\ln X_7 + 0.05481\ln X_8$	0.3017	0.0910	0.4019	5.257
(62)	$\ln Y = 0.02630 + 0.13925X_1 - 0.00310X_1X_2 + 0.00620X_1X_5$	0.3243	0.1052	0.3978	8.269
(63)	$\ln(Y/X_1) = -6.99917 - 0.80568X_1 + 0.05150X_3 - 0.05419X_7 + 0.19565X_1^2 - 0.00055X_1^2X_2 - 0.00499X_1^2X_3$	0.8621	0.7432	0.4040	100.325
(64)	$\log_{10}(Y/X_1) = -2.67225 - 0.28637X_1 + 0.01961X_1^2 + 0.00588(X_3-X_2)$	0.8320	0.6922	0.1907	158.191

[a] Standard Error of Estimate

estimate as 0.4040. The 74.32 percent of variation, in wet pavement traffic accidents at two-way bridge approaches, can be explained by six independent variables.

The comparison of dry and wet pavement accident models indicated that the relationship of wet pavement accidents and roadway elements is stronger than that for dry pavement conditions. The major independent variables in the case of wet pavement were volume of traffic, width of bridge, width of approaching pavement, and height of guard rail. The significant variables included in the case of dry pavement conditions were volume of traffic, percentage of grade, degree of curvature, and the cross-product of width of bridge and width of approaching pavement.

The linear regression analysis was also applied to four-lane roadways in studying the dry and wet pavement conditions. The sample sizes were 106 and 76 for dry and wet pavement conditions, respectively. The equations obtained are shown in Tables 32 and 33.

The best predictor model for dry pavement conditions was identified as model number 68 with the coefficient of multiple correlation of 72.94 percent. The variation in dry pavement accidents at one-way bridge approaches was explained by five variables. The coefficient of multiple

TABLE 32

DRY PAVEMENT TRAFFIC ACCIDENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(65)	$Y = 1.89439 + 0.00010X_1 + 0.00094X_8 - 0.02450X_9$	0.6375	0.4064	1.8154	23.275
(66)	$\ln Y = - 4.49478 + 0.44429 \ln X_1 + 0.17596 \ln X_8$	0.6032	0.3638	0.5822	29.455
(67)	$\ln Y = - 0.00214 + 0.00076X_1X_2 + 0.00003X_1X_8$	0.6996	0.4894	0.5215	49.370
(68)	$\ln(Y/X_1) = -7.86826 - 0.12860X_1 + 0.04457X_5 + 0.00041X_8 + 0.00002X_1^2X_2 + 0.0007X_1^2X_3$	0.7294	0.5320	0.5592	32.185
(69)	$\log_{10}(Y/X_1) = - 3.56220 - 0.05621X_1 + 0.01934X_5 + 0.00019X_8 + 0.00103X_1^2 + 0.00020X_2X_3$	0.7171	0.5142	0.2385	21.172

[a] Standard Error of Estimate

TABLE 33

WET PAVEMENT TRAFFIC ACCIDENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(70)	$\ln Y = -2.80222 + 0.35732 \ln X_1$	0.4170	0.1739	0.6112	15.577
(71)	$\ln Y = 0.14555 + 0.00073 X_1 X_2$	0.5115	0.2616	0.5770	26.215
(72)	$\ln(Y/X_1) = -7.61251 - 0.13976 X_1 - 0.00410 X_1^2 + 0.00003 X_1^2 X_2 + 0.00024 X_1^2 X_3 + 0.00004 X_4 X_5 + 0.03743 X_5 X_6$	0.7416	0.5500	0.5505	14.056
(73)	$\log_{10}(Y/X_1) = -3.51768 - 0.06032 X_1^2 + 0.00150 X_1^2 + 0.00024 X_2 X_3 + 0.00000 X_4 X_5 + 0.00006 X_4 X_6$	0.7337	0.5383	0.2404	16.324

[a] Standard Error of Estimate

determination was 53.20 percent with standard error of estimate as 0.5592. The significant variables were volume of traffic, width of bridge, width of approaching pavement, length of bridge, and percentage of grade.

In studying wet pavement conditions (four-lane roadway), the best predictor equation was identified as model number 72. The coefficient of multiple correlation was 74.16 with standard error of estimate of 0.5505. The variables explained 55.00 percent of variation in wet pavement traffic accidents at one-way bridge approaches. The significant variables in the model were volume of traffic, width of bridge, width of approaching pavement, cross product of grade and sight distance, and degree of curvature. It is interesting to note that the stopping sight distance has become a significant variable in wet pavement conditions but it was not significant to be included in dry pavement conditions reviewed above.

Single and Multiple Car Traffic Accidents

In an effort to delineate the effects of roadway elements on single and multiple car traffic accidents, the data were further classified according to single car and multiple car accidents. The sample sizes analyzed by the linear regression procedure were 459 and 245 for single car and multiple car traffic accidents, respectively. The equations derived for two-way bridges on two-lane highways are shown in Tables 34 and 35.

The examination of the statistics indicated that the highest coefficient of multiple correlation resulted from single car data in equation number 75. This coefficient was 81.28 percent with the standard error of estimate as 0.1941. The coefficient of multiple determination computed for this class was 66.07 percent, and the variation in single car traffic accidents was explained by three variables. The significant variables in the equation were volume of traffic, percentage of grade, and cross-product of bridge width and square of traffic volume. In contrast, multiple car traffic accidents showed a coefficient of multiple correlation lower than single car accidents, but the difference was not great. The coefficient of multiple determination was 58.32 percent with three significant variables in equation number 78. The variables included in the equation were volume of traffic, degree of curvature, and cross-product of width of bridge and width of approaching pavement.

In the analyses of four-lane highways (one-way bridge approaches) sample sizes of 111 as single car accidents

TABLE 34

SINGLE CAR TRAFFIC ACCIDENT

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(74)	$Y = 1.71066 + 0.00014X_1 - 0.02612X_2 + 0.03860X_5$	0.2274	0.0517	1.4169	8.268
(75)	$\log_{10}(Y/X_1) = -2.72498 - 0.24046X_1 + 0.00532X_5 + 0.02216X_1^2 - 0.00031X_1^2X_2$	0.8128	0.6607	0.1941	220.992
(76)	$\ln Y = 0.85860 + 0.12855\ln X_1 - 0.23211\ln X_2 + 0.01066X_5$	0.2841	0.0807	0.3678	13.320

[a] Standard Error of Estimate

TABLE 35

MULTIPLE CAR TRAFFIC ACCIDENT

TWO-LANE ROADWAY

Number	Equation	R	R^2	SEE[a]	F-Ratio
(77)	$Y = 3.49642 + 0.00023X_1 - 0.03239X_2 - 0.08076X_3$	0.3534	0.1249	1.4268	11.462
(78)	$\log_{10}(Y/X_1) = - 2.70215 - 0.21876X_1 + 0.00732X_6 + 0.01279X_1^2 - 0.00015X_2X_3$	0.7637	0.5832	0.2351	83.970
(79)	$\ln Y = 1.43427 + 0.20856\ln X_1 - 0.42136\ln X_2$	0.3265	0.1066	0.4826	14.438

[a] Standard Error of Estimate

and 78 as multiple car accidents were used. Since the trend of the data in previous classifications indicated an excellent suitability of model type 14 (see Table 23) to the data obtained, the linear regression analysis was performed on this model only. The results obtined through this analysis are shown in Tables 36 and 37.

The investigation of the equations revealed that single car traffic accidents showed a coefficient of multiple correlation approximately 27 percent higher than multiple car traffic accidents. The variables included in the equation of single car traffic accidents explained 64.33 percent of variation in response variable. The coefficient of multiple correlation was 80.21 percent; and significant variables were volume of traffic, height of guard rail, cross-product of percentage of grade and degree of curvature, and cross-product of volume squared and width of approaching pavement. It is also interesting to note that multiple car traffic accidents at one-way bridge approaches is a function of volume of traffic only and no geometric factors are involved in the derived model.

Width of Bridge Classification

The traffic accidents at bridge approaches of two-

TABLE 36

SINGLE CAR TRAFFIC ACCIDENT

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(80)	$\log_{10}(Y/X_1) = -3.36427 - 0.05646X_1 + 0.00009X_8 + 0.00004X_1^2X_3 + 0.02131X_5X_6$	0.8021	0.6433	0.2300	47.802
(81)	$\ln Y = -2.06168 + 0.21932\ln X_1 - 0.45613\ln X_2 + 1.11683\ln X_3$	0.4126	0.1703	0.5277	7.319

[a] Standard Error of Estimate

TABLE 37

MULTIPLE CAR TRAFFIC ACCIDENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(82)	$\log_{10}(Y/X_1) = -\ 3.37540 - 0.05103X_1 + 0.00095X_1^2$	0.5365	0.2878	0.2784	15.156
(83)	$\ln Y = -\ 1.46430 + 0.56930\ln X_1 + 0.12198\ln X_8$	0.5828	0.3396	0.6560	19.288

[a] Standard Error of Estimate

lane roadways were classified according to width of bridges as: (a) bridges twenty-four feet wide or less, and (b) bridges over twenty-four feet wide. The sample size for classes (a) and (b) were 349 and 262 distinct locations, respectively. Tables 38 and 39 show the models derived by linear regression analysis.

The best predictor equations are model numbers 87 and 93 in each classification. The examination of these two equations indicates that the effect of geometrics of roadways on traffic accidents is pronounced at bridges with less than twenty-four feet in width. A similar equation, derived for the bridges with over twenty-four feet in width, showed that the geometrics of highway are not a critical factor in accident occurrence, and the volume of traffic is the predominant variable. For example, equation number 87, derived for group (a), indicates that the significant factors are volume of traffic, degree of curvature, cross-product of percentage of grade and degree of curvature, and the cross-product of volume squared and width of bridge. Comparison of group (a) and (b) reveals that the only geometric factor involved in the group (b) is in a form of cross-product of volume squared and width of bridge. Although equation number 93 in group (b) has a significantly high coefficient of

TABLE 38

WIDTH OF BRIDGE LESS THAN 24.0 FEET

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(84)	$Y = 3.73205 + 0.00063X_1 - 0.04490X_{11}$	0.4441	0.1972	2.1312	42.508
(85)	$Y = 5.14090 + 0.00067X_1 - 0.21792X_2 + 0.06346X_5$	0.4455	0.1985	2.1327	28.475
(86)	$\ln Y = 1.99240 + 0.25420\ln X_1 - 0.68630\ln X_2 + 0.02668\ln X_6 + 0.05610\ln X_9$	0.4033	0.1627	0.5124	16.709
(87)	$\log_{10}(Y/X_1) = -2.56042 - 0.37380X_1 + 0.01510X_6 + 0.08751X_1^2 - 0.00232X_1^2X_2 + 0.00135X_5X_6$	0.7244	0.5248	0.2200	75.768
(88)	$\log_{10}Y = 0.04511 + 0.00327X_5 + 0.00897X_6 + 0.00768X_7 + 0.00002X_8$	0.4526	0.2048	0.2169	22.154

TABLE 38--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(89)	$\log_{10}(Y/X_1) = -2.24739 - 0.35729X_1 - 0.01536X_2 + 0.03491X_1^2 + 0.00068X_3X_6 + 0.00177X_5X_6$	0.6949	0.4829	0.2296	64.053

[a] Standard Error of Estimate

TABLE 39

WIDTH OF BRIDGE GREATER THAN 24.0 FEET

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(90)	$Y = 1.15853 + 0.00008X_1$	0.2541	0.0646	0.917	17.952
(91)	$Y = 1.15853 + 0.00008X_1$	0.2541	0.0646	0.9172	17.952
(92)	$\ln Y = 1.69884 + 0.15000\ln X_1 - 0.50313\ln X_4$	0.3074	0.0945	0.4092	13.514
(93)	$\log_{10}(Y/X_1) = -2.79736 - 0.17046X_1 - 0.00812X_1^2 - 0.00008X_1^2X_2$	0.8144	0.6633	0.2194	169.418
(94)	$\log_{10}Y = 0.03906 + 0.02173X_1 - 0.00422X_1X_6 + 0.00001X_1X_8$	0.3145	0.0989	0.1776	9.441
(95)	$\log_{10}(Y/X_1) = -2.80844 - 0.16042X_1 + 0.00410X_1^2$	0.8115	0.6586	0.2205	249.782

[a] Standard Error of Estimate

multiple correlation, it should not be considered as a proper equation in studying the influence of roadway geometrics on accidents at bridge approaches.

The coefficient of multiple determination is 52.48 percent for group (a). The coefficient of multiple correlation is given as 72.44 percent with a standard error of estimate of 0.2200.

ADT Classification

The classification of traffic accidents according to ADT was considered for two-way bridges on two-lane roadways. The data were examined in detail and it was found that 1800 ADT would be a logical choice for grouping the data into two classes. This arrangement created in each group a balanced data bank which was representative of traffic accidents on roadway system in the State. Thus, group (a) consisted of traffic accidents at bridge approaches with ADT under 1800, and group (b) included the traffic accidents at bridge approaches which had a traffic volume of over 1800 vehicles per day. In comparing ADT's to the capacity (19) of rural highways, it was found that the majority of rural highways operate under capacity. In general, the volume of 1800 ADT is approximately twenty-five percent of operating capacity of rural highways in the State of Oklahoma. However, this percentage is higher for some major highways which are in the vicinity of urban areas and are operating between two major urban centers. The sample sizes resulting from this grouping were 321 and 290 distinct locations for class (a) and (b), respectively. The output of linear regression analysis on each group is shown in Tables 40 and 41.

A review of the equations shown in Tables 40 and 41 reveals that the coefficients of multiple determination of class (b) are generally lower than those of class (a). In other words, a better coefficient of multiple correlation has been obtained for roadways which carry a traffic volume under 1800 vehicles per day. Equation number 100 is the best predictor model for class (a) with the coefficient of multiple determination 63.01 percent and the standard error of estimate 0.1712. The significant variables are volume of traffic, cross-product of width of approaching pavement and percentage of grade, cross-product of stopping sight distance and degree of curvature, and the difference between pavement width and bridge width. The coefficient of multiple determination is 32.06 percent for class (b), indicating that 32.06 percent of variation in traffic accidents is explained by three variables. The standard error of estimate is

TABLE 40

TRAFFIC VOLUME LESS THAN 1800 ADT

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(96)	$Y = 1.73988 + 0.00011X_1 + 0.00014X_4 + 0.03590X_5 + 0.06446X_6 - 0.00980X_{11}$	0.3696	0.1366	0.8484	9.968
(97)	$Y = 1.39973 + 0.00013X_1 - 0.01914X_2 + 0.00014X_4 + 0.03342X_6 + 0.07577X_7$	0.3662	0.1341	0.8496	9.760
(98)	$\ln Y = 0.96045 + 0.11860 \ln X_1 - 0.24865 \ln X_2 + 0.01815 \ln X_5 + 0.2619 \ln X_6$	0.3550	0.1260	0.3757	11.394
(99)	$\log_{10}(Y/X_1) = -2.35160 - 0.48403X_1 - 0.00340X_2 + 0.00627X_5 + 0.01204X_6 + 0.03463X_1^2 + 0.00106X_5X_6$	0.7890	0.6225	0.1732	86.313

TABLE 40--Continued

Number	Equation	R	R^2	SEE	F-Ratio
(100)	$\log_{10}(Y/X_1) = -\ 2.42607 - 0.48737X_1 + 0.03490X_1^2 + 0.00035X_3X_5 + 0.00001X_4X_6 + 0.00375(X_3-X_2)$	0.7938	0.6301	0.1712	107.306

[a] Standard Error of Estimate

TABLE 41

TRAFFIC VOLUME GREATER THAN 1800 ADT

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(101)	$Y = 6.63661 + 0.00028X_1 - 0.00033X_4 - 0.06722X_{11}$	0.3895	0.1517	2.3753	17.053
(102)	$Y = 5.64943 + 0.00030X_1 - 0.04363X_2 - 0.14348X_3 - 0.00034X_4$	0.3055	0.0933	2.4601	7.332
(103)	$\ln Y = 4.79428 + 0.42906\ln X_1 - 0.41271\ln X_2 - 0.84628\ln X_3 - 0.13292\ln X_4$	0.3721	0.1384	0.5531	11.448
(104)	$\log_{10}(Y/X_1) = -2.98099 - 0.05072X_1 + 0.00987X_6 - 0.0023X_2X_3$	0.5662	0.3206	0.2421	33.615
(105)	$\log_{10}(Y/X_1) = -2.97425 - 0.05098X_1 - 0.00004X_4 + 0.01013X_6 - 0.00024X_2X_3$	0.5660	0.3204	0.2421	33.586

[a] Standard Error of Estimate

0.2421, and significant variables are volume of traffic, degree of curvature, and cross-product of width of bridge and width of approaching pavement.

Classification of Roadway Surfaces

In order to examine the effect of pavement surface type on traffic accidents, the data were classified according to and analyzed on the basis of bituminous and concrete pavements. On two-lane roadways the sample sizes of traffic accidents were 450 for bituminous and 161 for concrete surfacing. This grouping indicated that approximately thirty -six percent of traffic accidents (in this study) occurred on pavements with concrete wearing surfaces.

The inspection of the equations, shown in Tables 42 and 43, indicated that equation numbers 109 and 110 are potential models to be selected. Model number 110 is chosen for the class of bituminous surfaces. The coefficient of multiple determination is 60.76 percent with standard error of estimate as 0.2263. The significant variables are traffic volume, degree of curvature, cross-product of volume and bridge width, and cross-product of volume and pavement width. In contrast to bituminous pavements, equation number 115 is selected for concrete pavements with the coefficient

TABLE 42

TRAFFIC ACCIDENTS ON BITUMINOUS PAVEMENTS

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(106)	$Y = 1.42303 + 0.00009X_1 - 0.01443X_2 + 0.00012X_4 + 0.05417X_6$	0.2625	0.0689	1.0128	8.234
(107)	$\ln Y = 0.15195 + 0.00002X_1 - 0.00337X_2 + 0.00427X_5 + 0.00813X_6$	0.2510	0.0630	0.1900	7.479
(108)	$\ln Y = -0.17136 + 0.14073\ln X_1 - 0.19758\ln X_2 + 0.02034X_6$	0.2818	0.0794	0.4331	12.828
(109)	$\log_{10}(Y/X_1) = -2.87687 - 0.11076X_1 + 0.01448X_6$	0.7211	0.5200	0.2497	242.091
(110)	$\log_{10}(Y/X_1) = -2.75287 - 0.19218X_1 + 0.01107X_6 - 0.00011X_1^2X_2 + 0.00044X_1^2X_3$	0.7795	0.6076	0.2263	172.253

[a] Standard Error of Estimate

TABLE 43

TRAFFIC ACCIDENTS ON CONCRETE PAVEMENTS

TWO-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(111)	$Y = 7.60942 + 0.00105X_1 - 0.05884X_2 - 0.26816X_3 - 0.00116X_4 + 0.10402X_5$	0.5423	0.2941	2.7663	12.914
(112)	$\ln Y = 0.62546 + 0.00009X_1 - 0.00377X_2 - 0.02622X_3 - 0.00010X_4 + 0.01154X_6 + 0.00012X_8$	0.5895	0.3475	0.2264	13.671
(113)	$\ln Y = 0.55329 + 0.43008\ln X_1 - 0.45099\ln X_2 - 0.29007\ln X_4 + 0.02359X_6 - 0.03723X_9$	0.5427	0.2945	0.5403	12.944
(114)	$\log_{10}(Y/X_1) = -2.37276 - 0.05633X_1 - 0.00517X_2 - 0.02021X_3 - 0.00010X_4 + 0.01015X_6$	0.6273	0.3935	0.2612	20.116
(115)	$\log_{10}(Y/X_1) = -2.54948 - 0.31215X_1 + 0.01421X_6 + 0.04702X_1^2 - 0.00014X_1^2X_2 - 0.00048X_1^2X_3 - 0.00001X_3X_4$	0.7278	0.5296	0.2308	28.901

[a] Standard Error of Estimate

of multiple determination 52.96 percent and standard error of estimate as 0.2308. The significant variables in this model are identical to bituminous pavements with an additional term as cross-product of pavement width and sight distance. This indicates that the geometric elements which were identified by the analysis in each case (bituminous or concrete) have equal degree of importance in accident occurrence regardless of surface type.

Four-lane highways were also grouped into the two classes of bituminous and concrete pavements. The sample sizes used in this grouping were 108 bituminous and 28 concrete pavements. Equation numbers 117 and 125 (Tables 44 and 45) were identified to be the best predictor equations in bituminous and concrete class, respectively. The coefficient of multiple determination, for bituminous type, is 65.94 percent with standard error of estimate as 0.2770. The important variables are traffic volume, pavement width, and length of bridge. For concrete pavements, the coefficient of multiple determination is 78.34 percent with significant variables as: traffic volume, bridge width, approaching grade, and the cross-product of grade and degree of curvature.

TABLE 44

TRAFFIC ACCIDENTS ON BITUMINOUS PAVEMENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(116)	$Y = -2.72511 + 0.00017X_1 + 0.12455X_3 + 0.00091X_8$	0.6092	0.3711	2.9746	20.455
(117)	$\log_{10}Y = -0.29789 + 0.00002X_1 + 0.01404X_3 + 0.00011X_8$	0.6594	0.4348	0.2770	26.673
(118)	$\ln Y = -5.69481 + 0.60778\ln X_1 + 0.16035\ln x_8$	0.5967	0.3560	0.6775	29.027
(119)	$\log_{10}(Y/X_1) = -3.41381 - 0.00531X_1 - 0.07157X_7 + 0.00011X_8$	0.4169	0.1738	0.3025	7.293
(120)	$\log_{10}(Y/X_1) = -3.38471 - 0.04226X_1 + 0.02133X_5 + 0.00010X_8 + 0.00003X_1^2X_3$	0.5121	0.2622	0.2873	9.153

[a] Standard Error of Estimate

TABLE 45

TRAFFIC ACCIDENTS ON CONCRETE PAVEMENTS

FOUR-LANE ROADWAYS

Number	Equation	R	R^2	SEE[a]	F-Ratio
(121)	$Y = 3.33006 - 0.10299X_3 + 0.00140X_8$	0.5607	0.3144	0.5554	5.733
(122)	$\ln Y = 0.60385 - 0.02658X_3 + 0.00033X_8$	0.5617	0.3156	0.1351	5.763
(123)	$\ln Y = 0.12682 + 0.05476\ln X_5$	0.3400	0.1156	0.3467	3.398
(124)	$\log_{10}(Y/X_1) = -3.34320 - 0.05842X_1 + 0.04255X_5 + 0.06937X_6$	0.8667	0.7511	0.1566	24.141
(125)	$\log_{10}(Y/X_1) = -3.17845 - 0.05379X_1 - 0.00575X_2 + 0.03249X_5 + 0.03064X_5X_6$	0.8851	0.7834	0.1493	20.791

[a] Standard Error of Estimate

Summary of Best Selected Linear Models

The investigation of traffic accidents at bridge approaches consisted of the study of twenty potential models which are shown in Table 23. In order to obtain a good fit, the number of accidents was divided by traffic volume and then, the logarithim of this ratio was treated as a response variable. Furthermore, the traffic accident data in each roadway classification were stratified according to weather condition, light condition, single and multiple car involvement, width of bridge, volume of traffic, and type of pavement surface (bituminous or concrete). The influence of the geometric elements of roadway on traffic accidents was studied by analyzing these models under classifications mentioned above.

The analyses of the potential models resulted in one hundred and twenty-five equations which are presented in Tables 24 through 45. The plot of residuals for these equations, obtained by linear regression analyses, was examined in order to investigate the violation of the assumptions concerning residuals. During the process of this examination, there was no apparent reason, based on the data used, to believe that the assumptions regarding residuals were violated. Then, the models which satisfied the statistical tests of hypotheses at an acceptable level, and had reasonable parameters with potential applications were selected as the best predictor models. The summary of selected models in each class of roadway and under each classified conditions is shown in Tables 46 and 47.

It is imperative that the nature of utilization and limitations of linear models be recognized. First, the linear models (resulting from regression analysis) are ordinarily a good forecasting tool if the conditions under which the data were collected remain reasonably stable. Second, accurate estimates can be obtained within the range of the data used in derivation, however, the accuracy of estimates decreases as one departs from the range of the data. The independent variables used in this study had a wide range of variation. The data range of response variable was not as wide as independent variables. The major concentration of traffic accidents lies between one to three accidents per year. Although the range of one to three accidents is small considering a section of roadway, this range would be significant if it occurred at a particular bridge site. Consequently, the models derived will provide the most accurate estimates within the range of the data and reasonable estimates, depending on the purpose of utilization, outside the

TABLE 46

SUMMARY OF BEST SELECTED MODELS

TWO-LANE ROADWAYS

Classification	Equation	R	R^2	F-Ratio
Overall Data (13)	$\log_{10}(Y/X_1) = -2.76381 - 0.19527X_1 + 0.01177X_6 + 0.01516X_1^2 - 0.00020X_1^2X_3$	0.7405	0.5483	183.892
Daytime (39)	$\log_{10}(Y/X_1) = -2.63826 - 0.26336X_1 + 0.01298X_6 + 0.01789X_1^2 - 0.00016X_2X_3$	0.8033	0.6453	163.747
Night Time (44)	$\log_{10}(Y/X_1) = -2.59442 - 0.31732X_1 + 0.00003X_4 + 0.02154X_1^2 - 0.00011X_2X_3$	0.8605	0.7405	238.962
Dry Pavement (59)	$\log_{10}(Y/X_1) = -2.61542 - 0.27551X_1 + 0.00471X_5 + 0.00827X_6 + 0.01811X_1^2 - 0.00011X_2X_3$	0.8023	0.6437	169.849
Wet Pavement (63)	$\ln(Y/X_1) = -6.9917 - 0.80568X_1 + 0.05150X_3 - 0.05414X_7 + 0.19565X_1^2 - 0.00055X_1^2X_2 - 0.00499X_1^2X_3$	0.8621	0.7432	100.325

TABLE 46--Continued

Classification	Equation	R	R^2	F-Ratio
Single Car (75)	$\log_{10}(Y/X_1) = -2.72498 - 0.24046X_1 + 0.00532X_5 + 0.02216X_1^2 - 0.00031X_1^2X_2$	0.8128	0.6607	220.992
Multiple Car (78)	$\log_{10}(Y/X_1) = -2.70215 - 0.21876X_1 + 0.00732X_6 + 0.01279X_1^2 - 0.00015X_2X_3$	0.7637	0.5832	83.970
WB[a] $<$ 24 (87)	$\log_{10}(Y/X_1) = -2.56042 - 0.37380X_1 + 0.01510X_6 + 0.08751X_1^2 - 0.00232X_1^2X_2 + 0.00135X_5X_6$	0.7244	0.5248	75.768
WB[a] $>$ 24 (93)	$\log_{10}(Y/X_1) = -2.79736 - 0.17046X_1 - 0.00812X_1^2 - 0.00008X_1^2X_2$	0.8144	0.6633	169.418
1800 ADT or Less (100)	$\log_{10}(Y/X_1) = -2.42607 - 0.48737X_1 + 0.03490X_1^2 + 0.00035X_3X_5 + 0.00001X_4X_6 + 0.00375(X_3 - X_2)$	0.7938	0.6301	107.306

TABLE 46--Continued

Classification	Equation	R	R^2	F-Ratio
Over 1800 ADT (104)	$\log_{10}(Y/X_1) = -2.98099 - 0.05072X_1 + 0.00987X_6 - 0.00023X_2X_3$	0.5662	0.3206	33.615
Bituminuous Pavements (110)	$\log_{10}(Y/X_1) = -2.75287 - 0.19218X_1 + 0.01107X_6 - 0.00011X_1^2X_2 + 0.00044X_1^2X_3$	0.7795	0.6076	172.253
Concrete Pavements (115)	$\log_{10}(Y/X_1) = -2.54948 - 0.31215X_1 + 0.01421X_6 + 0.04702X_1^2 - 0.00014X_1^2X_2 - 0.00048X_1^2X_3 - 0.00001X_3X_4$	0.7278	0.5296	28.901

[a]Width of Bridge.

TABLE 47

SUMMARY OF BEST SELECTED MODELS

FOUR-LANE ROADWAYS

Classification	Equation	R	R^2	F-Ratio
Overall Data (21)	$\ln Y = -0.11629 + 0.00004X_1^{a} + 0.02669X_3 + 0.03992X_5 + 0.00028X_8 - 0.00846X_9$	0.7152	0.5116	27.231
Daytime (49)	$\log_{10}(Y/X_1) = -3.63452 - 0.04589X_1 + 0.00011X_8 + 0.00087X_1^2 + 0.00019X_2X_3 + 0.00004X_4X_6$	0.6193	0.3835	12.689
Night Time (54)	$\log_{10}(Y/X_1) = -3.31152 - 0.06813X_1 + 0.06104X_6 + 0.00110X_1^2$	0.7995	0.6392	49.012
Dry Pavement (68)	$\ln(Y/X_1) = -7.86826 - 0.12860X_1 + 0.04457X_5 + 0.00041X_8 + 0.00002X_1^2X_2 + 0.0007X_1^2X_3$	0.7294	0.5320	32.185
Wet Pavement (72)	$\ln(Y/X_1) = -7.61251 - 0.13976X_1 - 0.00410X_1^2 + 0.00003X_1^2X_2 + 0.00024X_1^2X_3 + 0.00004X_4X_5 + 0.03743X_5X_6$	0.7416	0.5500	14.056

TABLE 47--Continued

Classification	Equation	R	R^2	F-Ratio
Single Car (80)	$\ln(Y/X_1) = -3.36427 - 0.05646X_1 + 0.00009X_8 + 0.00004X_1^2X_3 + 0.02131X_5X_6$	0.8021	0.6433	47.802
Multiple Car (83)	$\ln Y = -1.46430 + 0.56930\ln X_1 + 0.12198\ln X_8$	0.5828	0.3346	19.288
Bituminous Pavements (117)	$\log Y_{10} = -0.29789 + 0.00002X_1 + 0.01404X_3 + 0.00011X_8$	0.6594	0.4348	26.673
Concrete Pavements (125)	$\log_{10}(Y/X_1) = -3.17845 - 0.05379X_1 - 0.00575X_2 + 0.03249X_5 + 0.03064X_5X_6$	0.8851	0.7834	20.791

[a]The volume of traffic is expressed in terms of vehicles per day.

range of data. Since roadways are generally designed for a service life of twenty years, the estimating accuracy of the derived models should be good for at least several years. Any improvement in roadway elements is not expected to be significant in the first several years to require revision of the models. The range of data for the derived models are shown in Appendix B.

Accuracy and Limitations of Data

One of the most important factors in any type of research work is the degree of accuracy of the data used in developing a method of solution. Unfortunately, the biases in traffic accident records and related data are old known facts. For example, the collection of traffic records starts with the highway patrol investigating the scene of the accident. The multiple responsibilities of the policeman in the field prevent him from collecting detailed accident information regarding the causative factors. The general statements such as "too fast for the condition" or "driver lost control of car" are frequently found in the accident records. These and similar types of statements which are used for legal and insurance purposes overcast the inquiry for the real causes of traffic accidents. The second bias results from processing the data and transferring them onto the magnatic tape for automatic machine output. In dealing with massive data such as traffic accidents, every detail of the data or output can not be examined thoroughly. Therefore, the approximations which are made throughout the processing and merging of roadway factors with accident information are tolerated. The third bias stems from matching a particular bridge accident with the corresponding location of the bridge. The approximations and assumptions were essential in order to assign a proper bridge accident to its corresponding bridge location. The last bias is the human error which results in measuring the geometric elements of the roadways. The geometric parameters were taken from bridge and roadway plan-profile maps. The approximations were made where uncertainities were involved and the best values of geometric parameters were recorded.

Consequently, the models derived in this study are the best and the most accurate equations under the prevailing accuracy and limitations of the data. It must be noted that the word "approximation" mentioned above does not refer to "guess work", but rather to the engineering evaluation and practices employed in measuring and selecting the values of parameters.

Applications of Derived Models

In order to reduce traffic accidents at bridge approaches, often the spot improvement type of project is necessary to correct the existing deficiencies. The bridges are ordinarily located over drainage passages where the grades and curvatures are predominant. Thus, in many cases, extensive construction can be avoided and the corrections can be limited to removal of deficient roadway elements at the bridge vicinity only.

The models developed undoubtedly have many applications in practice as well as in theory in the field of highway and traffic engineering. Several important applications of the equations are cited for illustrative purposes in the following sections.

As a Decision Making Model

The models developed can be used as decision making models in highway engineering. This technique will enable practicing engineers to examine the feasibility of construction as a spot improvement based on the direct cost of accidents. The following example outlines the method.

Example 1:

Assume that the following data are given for an existing bridge location on a two-lane facility.

Accidents for a three-year period:

1 single car fatality

3 multiple car injuries

3 single car property damage

ADT (X_1) = 2,000 vehicles per day

X_2 = 20.0 feet

X_3 = 20.0 feet

X_4 = 300 feet

X_5 = -4.3 percent

X_6 = 6.0 degrees

X_7 = 4.0 feet

A new bridge structure and alignment is proposed for this location, and the data are given below.

ADT (X_1) = 4,000 vehicles per day

X_2 = 30.0 feet

X_3 = 24.0 feet

X_4 = 1,500 feet

X_5 = -1.0 percent

X_6 = 1.0 degree

X_7 = 4.0 feet

If the cost of construction for the project is estimated as \$60,000, investigate the economic feasibility of this construction project in reducing accidents at this location.

Solution:

Considering the equation derived for wet pavement accidents (see Table 46, equation 63), the expected value of accident at this location can be computed as follows:

$$\ln\left(\frac{Y_{63}}{X_1}\right) = -6.99170 - 0.80568(4) + 0.05150(24)$$
$$-0.05419(4) + 16(0.19565 - 0.00055(30)$$
$$-0.00499(24)) = -8.24494$$

$$\ln Y_{63} - \ln X_1 = -8.24494$$
$$\ln Y_{63} = \ln X_1 - 8.24494$$
$$= \ln(4000) - 8.24494$$
$$= +8.29404 - 8.24494 = 0.04910$$

$$Y_{63} = \text{Antiln}(0.04910) \qquad Y_{63} = 1.05$$

$$\text{Percent Reduction} = \left(\frac{2.33 - 1.05}{2.33}\right) 100 = 55.0$$

The average cost per accident using Appendix C can be calculated as:

$$\text{Average Cost/Accident} = \frac{6,200+3(3,600)+3(300)}{7} = \$2,472$$

The accidents saved during the service life of the project and unit cost per accident saved are computed below:

$$\text{Accidents Reduced} = (2.33)\ (0.55)\ (20) = 26$$

$$\text{Cost/Accident Reduced} = \frac{60,000}{26} = \$2,300$$

Comparison of these unit costs, 2,472 versus 2,300, indicates that the construction of this project based on direct cost of accident is justified. It is suggested that, in practice, all alternative formulas be computed, and the cost analysis should be based on the minimum reduction in accidents.

As a Priority Scheduling Technique

The annual budget allocated for construction and improvement of deficient locations is ordinarily limited. As a result, some of the construction projects would be postponed to future years. In order to construct the most needed projects as early as possible, a priority program is necessary. The method which will be used in establishing a priority program is based on the least cost per expected accident reduced as developed by the California Division of Highways (46). The example below will illustrate the use of the models developed in this study and the method of least cost in establishing a priority schedule for spot improvement of bridge locations.

<u>Example 2</u>:

Assume that there are n bridge locations which are in need of improvement. If a limited amount of funds is available during a certain fiscal year, establish a priority schedule of construction by use of least cost method based on the expected number of accidents reduced per year.

Solution:

The solution of this problem can be best demonstrated by outlining the method in the following steps:

1. Estimate the expected number of accidents per year using the derived models of this study, and compute the reduction in accidents as: $a_i = Y_i - \hat{Y}_i$, $(i=1,2, \ldots, n)$, where Y_i is the existing average accidents per year, and $\hat{Y}_i$ is the expected number of accidents per year after improvement.

2. Calculate the expected cost per accident reduced as outlined in example 1. Denote this cost as C_{rn}.

3. List the costs (C_{rn}) in ascending order and also tabulate all other corresponding information as shown in Table 48.

Comparison of available funds and cumulative cost of the projects is made in order to determine the cut-off point. This point can be established where the cumulative cost of projects is equal to the funds available for a particular fiscal year. Then, the number of projects which lie

TABLE 48

PRIORITY SCHEDULING TECHNIQUE -- EXAMPLE 2

Cost/Accident Reduced	Cost/Project	Cumulative Cost of Project	a_i	Cumulative a_is
c_{r1}	C_{p1}	C_{p1}	a_1	a_1
c_{r2}	C_{p3}	$C_{p1} + C_{p2}$	a_2	$a_1 + a_2$
c_{r3}	C_{p3}	$C_{p1} + C_{p2} + C_{p3}$	a_3	$a_1 + a_2 + a_3$
.	.	. . .	.	. . .
.	.	. . .	.	. . .
.	.	. . .	.	. . .
c_{rn}	C_{pn}	$C_{p1} + C_{p2} + \ldots + C_{pn}$	a_n	$a_1 + a_2 + \ldots + a_n$

above this line can be planned for construction during first fiscal year. This process should be repeated for each fiscal year.

In order to illustrate the method presented above by a hypothetical example, assume that a state government has allocated $300,000 for spot improvement of various bridge locations for a particular fiscal year. If all the bridge sites were studied, and the plans for improvement were developed, the estimated cost for each site can be given as:

Bridge Site A:	$ 50,000
B:	$100,000
C:	$ 30,000
D:	$ 80,000
E:	$ 45,000
F:	$ 70,000
. . . . . .	
. . . . . .	

An "n" number of bridge sites are considered for possible improvement, and these require an investment of money as indicated above. The values of a_i and c_{rn} are obtained as explained above and once the total project costs (C_{pn}) are calculated, the improvements to be undertaken in that particular fiscal year are determined. The resulting priority schedule is shown in Table 49. In improving these projects, the cut-off point lies between $245,000 and $325,000. Therefore, the government can efficiently implement the first four projects at a cost of $245,000 during the first fiscal year. The examination of cumulative cost shows $55,000 surplus after the improvements are made, and this sum of money should be utilized in the next year's program from the view point of engineering analysis. Further inspection indicates that bridge site A requires only $50,000 for construction, and it can easily be financed by surplus money. The surplus money can either be used to construct bridge site A or it can be used in the next year's budget. This remains to be a political decision that should be made by the political governing body. The remaining projects should be re-scheduled for the next year.

As a Means of Comparative Study

It is sometimes desirable in practice to study a bridge location regarding traffic accidents under different environmental conditions. The equations developed will enable traffic engineers to compute the expected value of

TABLE 49

PRIORITY SCHEDULING TECHNIQUE--EXAMPLE 2-A NUMERICAL ANALYSIS

Bridge Site Number	Project Cost (C_{pn}),$	a_i [a]	Cumulative a_i	Cost/Accident[a] Reduced (C_{rn}),$	Priority	Cumulative Cost of Projects,$
C	30,000	1.50	1.50	1,000	1	30,000
E	45,000	1.30	2.80	1,700	2	75,000
F	70,000	1.60	4.40	2,200	3	145,000
B	100,000	2.00	6.40	2,500	4	245,000
D	80,000	1.00	7.40	4,000	5	325,000
A	50,000	0.50	7.90	5,000	6	325,000
.	. . .	. .	. .	. . .	.	. . .
.	. . .	. .	. .	. . .	.	. . .
.	. . .	. .	. .	. . .	.	. . .

[a] Calculated as in Example 1.

traffic accidents, based on geometrics of roadway, at the vicinity of bridge approaches. This is illustrated in the example below.

<u>Example 3</u>:

Assume that a bridge location, with the data given below, has one accident in the daytime and two accidents at nighttime during a year. Does this imply that darkness is a significant factor in accident occurrence at this location?

Bridge data are: ADT = 1,300, bridge width = 20, pavement width = 24, sight distance = 300, grade = + 3.2, and degree of curvature = 5.0.

Using formulas 39 and 44, Table 46, the expected value of accident for each case is:

Day: $\log_{10}(Y/X_1) = -2.63826 - 0.2633(1.3) + 0.01298(5.0) + 0.01789(1.3)^2 - 0.00016(20)(24)$

$$= -3.0000$$

$$\log_{10}Y = \log_{10}X_1 - 3.0000 = 3.1139 - 3.0000$$

$$= 0.1139 \quad \text{or} \quad Y = 1.4$$

Night: $\log_{10}(Y/X_1) = -2.59442 - 0.31732(1.3) + 0.00003(300) + 0.02154(1.3)^2 - 0.00011(20)(24)$

$$= -3.0100$$

$$\log_{10}Y = \log_{10}X_1 - 3.0100 = 3.1139 - 3.0100$$

$$= 0.1039 \quad \text{or} \quad Y = 1.3$$

$$\text{Chi-square} = (1.0 - 1.4)^2/1.4 + (2.0 - 1.3)^2/1.3$$

$$= 0.489$$

The tables for the Chi-square distribution with one degree of freedom and at the 5 percent significant level indicate a value of Chi-square as 3.84. Therefore, the test does not substantiate a significant difference and there is no reason to assume that darkness is a factor in accident occurrence at this location. If the observed number of the accidents at nighttime were 3.6 (instead of 2.0), the difference would have been significant at 5 percent level of significance. In this case some corrective measures are necessary to improve the situation at night.

CHAPTER VII

SUMMARY AND CONCLUSIONS

The effects of the geometric elements of roadway on traffic accidents at bridge approaches were investigated during the conduct of this research. The most significant geometric elements affecting traffic accidents were identified, and one hundred and twenty-five linear models were derived to predict traffic accidents under different environmental conditions. The best predictor models were chosen based on statistical tests of hypotheses namely examination of the coefficient of multiple determination and F-ratio at 0.05 level of significance.

Furthermore, the applications of the equations selected were illustrated by example problems and their practical aspects were pointed out. The most outstanding implications of this study can be utilized in the administrative decision making process, planning of spot improvement projects, and traffic engineering studies.

This study provided an insight concerning the role of geometric elements on traffic accidents at bridge approaches. The specific conclusions derived from the study are presented below.

1. The traffic accidents at bridge approaches did not have a linear relationship with geometric elements of roadway. This non-linear relationship was expressed by logarithmic transformations very well.

2. Average daily traffic volume was one of the most significant variables explaining the variation in traffic accidents in all models derived.

3. A combination of roadway elements is the most likely factor contributing to traffic accidents at bridge approaches, and the narrow bridge width (lower than 24 feet) is not the only factor causing the traffic accidents.

4. The models developed for nighttime and wet pavement accidents on two-way bridge approaches have the highest coefficient of multiple determination. The accident prone elements of roadways, at night hours, were identified as sight distance, width of bridge, and width of approaching pavement. In contrast, the roadway elements in case of wet pavement conditions were width of approaching pavement, width of bridge, and height of the guard rails.

5. For single car accidents the most significantly contributory roadway elements were percentage of approaching grade and width of the bridge. In comparison, the accident prone roadway elements in case of multiple car involvement were degree of curvature and width of the approaching pavement.

6. When sites with over twenty-four feet bridge width were studied, the width of bridge was identified as a contributing element to traffic accidents. The accident prone elements for bridges of less than twenty-four feet of width were the width of bridge, degree of curvature, and percentage of approaching grade.

7. Traffic accidents in the 0-1800 ADT class (1800 ADT being the breaking point) showed a highly significant coefficient of multiple correlation. The roadway elements for this group were a combination of width of bridge and of pavement, percentage of approaching grade, degree of curvature, and stopping sight distance. For the class of over 1800 ADT, the roadway elements were width of bridge and of pavement, and degree of curvature.

The 1800 ADT represent approximately twenty-five percent of operating capacity of rural highway (in the State of Oklahoma) and one-half of the sample size in the study was under this range. The examination of traffic volume showed that most of the rural highways are operating under capacity, and the capacity varies widely from one location to another.

8. The stratification was considered according to passenger cars and truck traffic accidents but the available data indicated a lack of sufficient number of truck traffic accidents and therefore the analysis was not justified.

9. Initially, the "speed before contact" and "speed change" were considered to be included in the analysis as independent variables. Since "speed before contact" had random characteristics and multicollinearity with other variables, it was excluded from computation in developing the final models. The lack of data concerning "speed change" variable prevented inclusion of this variable in the analysis.

10. Although simple models were obtained, non-linearity of relationship between traffic accidents and roadway geometrics was also evident for four-lane roadways.

11. The average daily traffic volume was one of the most significant variables contributing to accidents on one-way bridge approaches.

12. The degree of curvature combined with traffic volume are only contributing factors to traffic accidents at

night hours on four-lane roadways. This is an interesting result because the degree of curvature used on four-lane roadways is much smaller than that on two-lane roadways. The accident prone elements for wet pavement conditions were identified, by analysis, as width of bridge and of pavement, combination of sight distance and percentage of approaching grade, and a combination of degree of curvature and percentage of approaching grade.

13. For single car traffic accidents, the accident prone geometrics are length of bridge, width of pavement, and a combination of degree of curvature and percentage of approaching grade. An interesting result was that the multiple car involvement was only a function of traffic volume and length of bridge on four-lane highways.

14. Identical models were developed when the data were classified according to surface type (bituminous or concrete) of two-lane roadways. The common variables were traffic volume, cross-product of volume and bridge width, and cross-product of volume and pavement width. The model developed for concrete surface type had an additional term as cross-product of grade and degree of curvature.

The models developed for four-lane roadways indicated different equations for bituminous and concrete type surfaces. The significant variables identified for bituminous pavements were traffic volume, pavement width, and length of bridge. In comparison, important variables for concrete type of pavements were traffic volume, bridge width, grade, and cross-product of grade and degree of curvature. It appears that the grade is strongly related to accidents on four-lane roadways with concrete surfacing.

Recommendations for Future Studies

This research provided an insight to the nature of bridge accidents, and the role of geometric elements of roadways as contributory parameters to traffic accidents at bridge approaches. Thus, further research is suggested in the following areas.

First, during the course of this study it was found that an immediate need exists for collection of accurate data concerning geometric elements and traffic accidents at bridge approaches. A more precise methodology should be employed in determining the exact location of the accident relative to a particular bridge site.

The second proposal concerns the continuity concept in roadway design. The study indicated that many traffic accidents occurred at bridges with over twenty-four feet of width. This simply indicates that the discontinuity in roadway is a major factor in contributing to accidents. This discontinuity ordinarily appears when a roadway with twenty-four feet width and eight to ten feet shoulders reduces to a bridge passage of twenty-four feet (or under) width and about four feet high guard rails. It is suggested that a study be undertaken concerning the human factors which are contributing to traffic accidents at bridge approaches and to determine the association of roadway discontinuity and the human factors involved.

Third, the results of this study showed that the problem of traffic accidents at bridge approaches is the combination of several deficient roadway elements. Thus, a modification of design standards for bridge approaches is necessary. A successful study of this type will require extension of the present study and the collection of the necessary information on a regional basis to cover wider areas.

BIBLIOGRAPHY

1. Summary Description and Major Findings on the Study of Auto Accident Victims, Department of Transportation, Washington, D. C., 1970.

2. Rosenfield, Harry H., "President Sends Transportation Message to the Congress," Traffic Safety, April 1966, p. 8.

3. McGlade, Frank, "Traffic Accident Research: Review and Prognosis," Traffic Quarterly, Eno Foundation, Saugatuck, Connecticut, October 1962, pp. 559-560.

4. Little, Arthur D., The State of the Art of Traffic Safety, Praeger Publishers, New York, 1970, pp. 1-127.

5. Tarrants, William E., "Current Research in Driver Education," Highway Research Board Special Report 107, Washington, 1970, pp. 34-41.

6. Winfrey, Robley, Economic Analysis for Highways, International Textbook Company, Scranton, Pennsylvania, 1969, p. 365.

7. Minahan, Daniel J., The Highway Spot Improvement Program--A Critical Review, Highway Safety Research Institute, The University of Michigan, Ann Arbor, Michigan, 1970.

8. Report of Annual Accidents Records, Traffic Engineering Division, Oklahoma Department of Highways, Oklahoma City, 1968.

9. Report of Annual Accident Records, Traffic Engineering Division, Oklahoma Department of Highways, Oklahoma City, 1969.

10. Traffic Control and Roadway Elements--Their Relationship to Highway Safety, Automotive Safety Foundation, Saugatuck, Connecticut, 1963, pp. 3-70.

11. Schoppert, David W., "Predicting Traffic Accidents From Roadway Elements of Rural Two-Lane Highways with Gravel Shoulders," Highway Research Board Bulletin 158, Washington, D.C., 1957.

12. Head, J. A., "Predicting Traffic Accidents from Roadway Elements on Urban Extensions of State Highways," Highway Research Board Bulletin 208, Washington, D. C., 1958.

13. Versace, John, "Factor Analysis of Roadway and Accident Data," Highway Research Board Bulletin 240, Washington, D. C., 1960.

14. Raff, Morton S., "Interstate Highway Accident Study," Public Roads, Volume 27, No. 8, Washington, D.C., June 1953, pp. 170-186.

15. Cirillo, Julie A.; Dietz, Stephen K.; and Beaty, Richard L., Analysis and Modeling of Relationships Between Accidents and Geometrics and Traffic Characteristics of the Interstate System, Office of Research and Development, Traffic Systems Division, Bureau of Public Roads, Washington, D. C., August 1969, p.1.

16. Kihlbert, Jaakko K., and Tharp, K.J., "Accident Rates as Related to Design Elements of Rural Highways," Highway Research Board NCHRP No. 47, Washington, D.C., 1968, pp. 21-22.

17. Mohr, Harvey W., "Results of Speed Zoning on Rural Highways," Highway Research Board Proceedings, Washington, D.C., 1954, pp. 420-446.

18. Hoch, Irving, "Accident Experience: Comparing Expressways and Arterials," Traffic Quarterly, Volume 14, No. 3, Saugatuck, Connecticut, July 1960, pp. 340-362.

19. A Policy on Geometric Design of Rural Highways, American Association of State Highway Officials, Washington, D.C., 1965, pp. 225-261.

20. Belmont, D.M., "Effect of Average Speed and Volume on Motor Vehicle Accidents on Two-lane Tangents," Highway Research Board Proceedings, Volume 32, Washington, D.C., 1953, p. 383.

21. Belmont, D.M., "Accidents Versus Width of Paved Shoulders on California Two-lane Tangents--1951 and 1952," Highway Research Board Bulletin 117, Washington, D.C., 1956, pp. 1-5.

22. Billion, C.E. and Stohner, Walker R., "A Detailed Study of Accidents as Related to Highway Shoulders in New York State," Highway Research Board Proceedings, Washington, D.C., 1957, pp. 497-508.

23. Gunnerson, H.E., "Iowa Narrow Bridge Accident Study," Highway Research Board Abstracts, Volume 31, No. 7, Washington, D.C., 1961, p. 23.

24. Mullins, B.K. and Keese, J.C., "Freeway Traffic Accident Analysis and Safety Study," Highway Research Board Bulletin 291, Washington, D.C., 1961.

25. Bowman, B.H., "Ohio Turnpike Accident Study," Traffic Engineering, Volume 28, No. 9, June 1958, pp. 23-26.

26. Glennon, John C., Evaluation of Stopping Sight Distance Design Criteria, Summary Report 134-3(S), Texas Transportation Institute, Texas A & M University, College Station, Texas, January 1970.

27. Woods, Donald L., "Vertical Alignment Effects on Safe Vehicle Operation," Texas Transportation Researchers, Volume 6, No. 2, Texas Transportation Institute, College Station, Texas, April 1970.

28. Matson, T.M.; Smith, W.S.; and Hurd, F.W., Traffic Engineering, McGraw-Hill Book Company, New York, 1955.

29. Investigator Collision Report Coding Guide, Oklahoma Department of Public Safety, Oklahoma City, 1970.

30. Sufficiency Rating Report and Need Study, Planning Division, Oklahoma Department of Highways, Oklahoma City, 1968.

31. Oklahoma Traffic Characteristics--A Factual Analysis of Traffic Volumes and Vehicle Classification Data, Planning Division, Oklahoma Department of Highways, Oklahoma City, 1968.

32. Oklahoma Traffic Characteristics--A Factual Analysis of Traffic Volumes and Vehicle Classification Data, Planning Division, Oklahoma Department of Highways, Oklahoma City, 1969.

33. Graybill, Franklin A., Introduction to Linear Statistical Models, Volume I, McGraw-Hill Book Company, Inc., New York, 1961.

34. Mood, Alexander M., and Graybill, Franklin A., Introduction to the Theory of Statistics, McGraw-Hill Book Company, Inc., New York, 1963.

35. Neville, Adam A., and Kennedy, John B., Basic Statistical Methods for Engineers and Scientists, International Textbook Company, Scranton, Pennsylvania, 1964.

36. Draper, N. R., and Smith, H., Applied Regression Analysis, John Wiley and Sons, Inc., New York, 1966.

37. Rao, C. R., Advanced Methods in Biometric Research, John Wiley and Sons, Inc., New York, 1952, pp. 237-351.

38. Tintner, Gerhard, Econometrics, John Wiley and Sons, Inc., New York, 1958, pp. 126-154.

39. Rao, C. R., Linear Statistical Inference and Its Applications, John Wiley and Sons, Inc., New York, 1965, pp. 435-510.

40. Anderson, T. W., An Introduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New York, 1958, pp. 126-154.

41. Dixon, W. J., Ed. Biomedical Computer Programs, University of California Publication in Automatic Computation, University of California, Los Angeles, 1970.

42. Efroymson, M. A., "Multiple Regression Analysis," Mathematical Methods for Digital Computers, Volume I, John Wiley and Sons, Inc., New York, 1960, pp. 191-204.

43. Johnston, J., Econometric Methods, McGraw-Hill Book Company, Inc., New York, 1963, pp. 9-138.

44. Ashton, Winifred D., The Theory of Road Traffic Flow, John Wiley and Sons, Inc., New York, 1966, pp. 148-168.

45. Gerlough, Daniel L., "Use of Poisson Distribution in Highway Traffic," Poisson and Traffic, The Eno Foundation for Highway Traffic Control, Saugatuck, Connecticut, 1955, pp. 1-42.

46. Smith, Richard N. and Tamburri, Thomas N., "Direct Costs of California State Highway Accidents," Highway Research Board Record 225, Washington, D. C., 1968.

47. Little, Arthur D., Cost-Effectiveness in Traffic Safety, Frederick A. Praeger Publishers, New York, 1968.

48. Kaiser, Harold, "Traffic Accidents on Highways Bridges," Proceedings of the Tenth Annual Ohio Highway Engineering Conference, The Ohio State University, Columbus, Ohio, 1956.

49. Schwar, Johannes F. and Puy-Hurate, Jose, Statistical Methods in Traffic Engineering, Department of Civil Engineering, The Ohio State University, Columbus, Ohio, 1967.

50. Brownlee, K. A., Statistical Theory and Methodology in Science and Engineering, John Wiley and Sons, Inc., New York, 1965.

51. Lehmann, E. L., Testing Statistical Hypotheses, John Wiley and Sons, Inc., New York, 1959.

52. Morrison, Donald F., Multivariate Statistical Methods, McGraw-Hill Book Company, New York, 1967.

53. Cirillo, Julie Anna, "Interstate System Accident Research-Study II," Highway Research Board Record 188, Washington, D. C., 1967.

54. Cribbins, P. D.; Arey, J. M.; and Donaldson, J. K., "Effects of Selected Roadway and Operational Characteristics on Accidents on Multilane Highways," Highway Research Board Record 188, Washington, D. C., 1967.

55. Cribbins, P. D.; Horn, J. W.; Beeson, F. W.; and Taylor, R. D., "Median Openings on Divided Highways: Their Effect on Accident Rates and Level of Service," Highway Research Board Record 188, Washington, D. C., 1967.

56. Surti, Vasant H., "The Relationships of Vehicle Classification and Geometric Characteristics to Peak Period Freeway Volumes," Highway Research Board Record 162, Washington, D. C., 1967.

57. Tarrants, William E., "Myths and Misconceptions in Traffic Safety," Highway Research Board News 31, Washington, D. C., 1968.

58. Mosher, Walter W., Jr., "The Highway Environment and Safety," Traffic Safety: A National Problem, The Eno Foundation for Highway Traffic Control, Saugatuck, Connecticut, 1967.

59. Parzen, Emanuel, Modern Probability Theory and Its Applications, John Wiley and Sons, Inc., New York, 1960, p. 6.

60. Clark, N. and Pretty, R. L., editors, Traffic Engineering Practice, Department of Civil Engineering, University of Melbourne, Parkville, Victoria, Australia, 1969.

APPENDIX A

TABLES OF DISCRIMINANT ANALYSIS

TABLE 1

STEPWISE DISCRIMINANT ANALYSIS
DRY AND WET PAVEMENT ACCIDENTS - TWO-LANE ROADWAYS

Variables	Dry Coefficients	Wet Coefficients	U Statistics	F-Ratio
X_1	- 0.00510	- 0.00040	0.9843	11.0174
X_3	4.56394	4.50296	0.9757	6.0358
X_8	- 0.00167	- 0.00160	0.9721	2.5594
Constants	-50.55064	-49.65207		
Sample Size	476	215		

TABLE 2

STEPWISE DISCRIMINANT ANALYSIS
DAY AND NIGHT ACCIDENTS - TWO-LANE ROADWAYS

Variables	Day Coefficients	Night Coefficients	U Statistics	F-Ratio
X_6	0.64659	0.54481	0.99096	6.41548
X_1	0.00059	0.00053	0.98458	5.49693
Y	0.18939	0.12889	0.98067	4.60481
Constants	-1.45661	-1.05289		
Sample Size	365	340		

TABLE 3

STEPWISE DISCRIMINANT ANALYSIS
SINGLE AND MULTIPLE CAR ACCIDENTS - TWO-LANE ROADWAYS

Variables	Single Car Coefficients	Multiple Car Coefficients	U Statistics	F-Ratio
X_1	0.00049	0.00066	0.97058	21.2811
X_6	0.16492	0.26116	0.95327	12.7256
X_9	0.00088	0.00123	0.94503	6.1029
Constants	-0.78191	-1.52770		
Sample Size	459	245		

APPENDIX B

THE RANGE OF DATA USED IN DEVELOPING THE MODELS

TABLE 1

THE RANGE OF DATA USED IN DEVELOPING THE MODELS
TWO-LANE ROADWAYS

Classification	Number of Accidents	ADT	Bridge Width	Pavement Width	Sight Distance	Approach-ing Grade	Degree of Curvature	Guard Rail Height	Length of Bridge	Number of Openings
Overall Data	1-9	125-12,800	15-32	18-24	150-3,500	-8.4-+7.8	0-22.5	1.0-6.0	20-3,700	0-11
Daytime	1-8	400-12,800	15-32	18-24	200-3,500	-7.0-+6.9	0-22.5	1.0-6.0	22-3,700	0-10
Nighttime	1-9	125-11,000	15-32	18-24	200-3,500	-8.4-+7.8	0-15.0	1.0-6.0	20-3,700	0-11
Dry Pavement	1-6	125-12,800	15-32	18-24	150-3,500	-7.0-+7.8	0-22.5	1.0-6.0	20-3,700	0-11
Wet Pavement	1-9	300-11,000	15-32	18-24	150-3,500	-8.4-+6.5	0-22.5	1.0-6.0	22-3,700	0-10
Single Car	1-6	175-11,000	15-32	18-24	150-3,500	-8.4-+7.8	0-22.5	1.5-6.0	22-3,700	0-11
Multiple Car	1-8	200-12,800	15-32	18-24	150-3,500	-7.8-+7.8	0-20.0	1.5-6.0	22-3,700	0-10
WB $<$ 24	1-7	125- 9,000	15-24	18-24	150-3,500	-7.8-+7.8	0-22.5	1.5-6.0	22-3,700	0-11
WB $>$ 24	1-9	200-12,800	24-32	18-24	300-3,500	-6.0-+6.5	0-14.0	1.0-6.0	20-3,672	0-10
1800 ADT or less	1-6	0- 1,800	15-32	18-24	150-3,300	-8.4-+7.0	0-22.5	1.0-6.0	22-3,700	0-10
Over 1800 ADT	1-9	1,800-12,800	18-32	18-24	250-3,500	-5.8-+7.8	0-15.5	1.0-6.0	20-3,700	0-11
Bituminous Surface	1-9	125-12,800	15-32	18-24	150-3,500	-8.4-+7.8	0-22.5	1.0-6.0	22-3,700	0-11
Concrete Surface	1-8	350- 8,900	15-32	18-24	200-3,300	-5.9-+6.9	0-15.0	1.5-6.0	20-2,272	0- 7

TABLE 2

THE RANGE OF DATA USED IN DEVELOPING THE MODELS
FOUR-LANE ROADWAYS

Classification	Number of Accidents	ADT	Bridge Width	Pavement Width	Sight Distance	Approach-ing Grade	Degree of Curvature	Guard Rail Height	Length of Bridge	Number of Openings
Overall Data	1-18	800-44,300	18-44	18-24	300-4,000	-5.0-+5.7	0-5.0	1.5-5.0	22-3,700	0-5
Daytime	1-12	2,000-44,300	19-44	18-24	300-4,000	-4.4-+5.7	0-5.0	1.5-4.5	22-3,700	0-5
Nighttime	1- 7	800-44,300	18-44	18-24	300-3,500	-5.0-+5.7	0-5.0	1.5-5.0	22-3,700	0-4
Dry Pavement	1-15	2,000-44,300	19-44	18-24	300-4,000	-5.0-+5.7	0-5.0	1.5-4.5	30-3,700	0-4
Wet Pavement	1-13	800-44,300	18-44	18-24	300-3,500	-4.4-+5.7	0-5.0	2.0-5.0	22-3,700	0-5
Single Car	1-10	2,000-44,300	20-44	20-24	375-3,500	-5.0-+5.7	0-5.0	1.5-4.0	22-3,700	0-5
Multiple Car	1-13	3,400-44,300	19-44	18-24	300-4,000	-4.4-+5.7	0-5.0	2.0-5.0	33-3,700	0-4
Bituminous Surface	1-18	2,000-44,300	18-44	18-24	300-3,200	-4.4-+5.7	0-5.0	1.5-5.0	22-3,700	0-5
Concrete Surface	1- 3	800-17,000	19-44	18-24	600-4,000	-5.0-+4.4	0-4.0	2.1-4.0	30-1,345	0-4

APPENDIX C

TABLES OF ACCIDENT COSTS

TABLE 1

AVERAGE COSTS PER REPORTED ACCIDENT

Type of Facility	Fatal Accidents	Injury Accidents	Property Damage	All
Rural	$9,700	$2,500	$500	$1,600
Urban	7,700	2,000	300	1,000
Total	$9,000	$2,200	$400	$1,200

Source: Direct Costs of California State Highway Accidents, by R. N. Smith and T. N. Tamburri, Highway Research Board, Record 225, Washington, D. C., 1968.

TABLE 2

AVERAGE COSTS PER REPORTED AND UNREPORTED ACCIDENTS

Type of Facility	Fatal Accidents	Injury Accidents	Property Damage	All
Rural	$9,700	$2,600	$800	$1,800
Urban	7,700	2,100	600	1,200
Total	$9,000	$2,300	$700	$1,400

Source: Direct Costs of California State Highway Accidents, by R. N. Smith and T. N. Tamburri, Highway Research Board, Record 225, Washington, D. C., 1968.

TABLE 3

AVERAGE COSTS PER REPORTED SINGLE-VEHICLE ACCIDENTS

Type of Facility	Fatal Accidents	Injury Accidents	Property Damage	All
Rural	$6,200	$1,600	$300	$1,100
Urban	4,600	1,000	200	600
Total	$5,600	$1,300	$220	$ 900

Source: Direct Costs of California State Highway Accidents, by R. N. Smith and T. N. Tamburri, Highway Research Board, Record 225, Washington, D. C., 1968.

TABLE 4

AVERAGE COSTS PER REPORTED MULTIPLE-VEHICLE ACCIDENTS

Type of Facility	Fatal Accidents	Injury Accidents	Property Damage	All
Rural	$13,600	$3,400	$600	$2,100
Urban	11,200	2,400	300	1,100
Total	$12,700	$2,700	$400	$1,300

Source: Direct Costs of California State Highway Accidents, by R. N. Smith and T. N. Tamburri, Highway Research Board, Record 225, Washington, D. C., 1968.