

THESES IN COMPREHENSIVE HEALTH PLANNING

AN ATTEMPTED CORRELATION OF HOME
ACCIDENT FATALITIES AND HOUSING QUALITY

by

Charles Thomas Male III

Thesis No. 2

A Thesis

Presented to the Faculty of the Graduate School

of Cornell University for the Degree of

Master of Regional Planning

June 1968



CENTER FOR HOUSING AND ENVIRONMENTAL STUDIES, DIVISION OF URBAN
STUDIES, 109 WEST SIBLEY, CORNELL UNIVERSITY, ITHACA, NEW YORK 14850

TABLE OF CONTENTS

CHAPTER	Page
I. THE PROBLEM AND THE ACCIDENT MODEL	1
Introduction	1
Accidents--The Problem	4
Accidents--The General Model	8
Human Failure--Individual Characteristics	10
Environmental Hazards	14
Conclusion	16
II. THE DATA AND THE ORIGINAL REGRESSION	17
Introduction	17
The Fatality Data	18
The Dilapidation Data	19
Populations and Death Rates	21
Home Accidents and the Binomial Model	22
Choice of Analytical Techniques	25
The Initial Regression	27
Problems with Use of Regression on Raw Data	28
Variation in County Population Sizes	28
Variations in Population Age Distribution	31
Conclusion	32
III. FURTHER MODIFICATIONS	33
Introduction	33
The Equivalent Population	34
Use of Equivalent Population with 16 Largest Counties	37
Aggregating the Data	39
Conclusion	43

CHAPTER	Page
IV. CONCLUSION	44
APPENDIX	50
BIBLIOGRAPHY	76

CHAPTER I
THE PROBLEM AND THE ACCIDENT MODEL

Introduction

Of all the subject areas with which environmental health planning is concerned, the area of housing as it relates to health is the one in which our knowledge is the most vague and our theories the least tested. At the same time, most of the persons connected with planning or concerned with urban problems in general seem to adhere to the idea that there is something wrong with run-down, "dilapidated" housing, housing that is in a relatively high state of disrepair. This idea is also widely held, and accepted even more uncritically, on the part of the popular press. Almost daily we see statistics indicating that a high percentage of dwelling units in this or that slum area are "substandard," with the tacit, widely accepted assumption that this is bad. Rarely is this underlying assumption questioned. The idea that dilapidated housing is unhealthy has become a more or less accepted doctrine.

Why do we consider so called "substandard" housing to be detrimental to health? Is it because this idea has been scientifically proven, or is it because run-down housing tends to offend our aesthetic sensibilities, thus leaving us with a "gut" feeling that something must be wrong with it? The second state of affairs is probably closer to the case. The official dogma, of course, is that dilapidated housing is a threat to the health and safety of the individual, and this society can legislate against it for the benefit of the public "health, safety, and general welfare." Nevertheless, many things have been accepted by the courts as detrimental to the public health, safety, and general welfare without the case really having been proven one way or the other, and "substandard" housing has been one of these. There has been very little, in the way of either conclusive statistical evidence or proven theory, shown to indicate that the commonly accepted idea has a basis in actual fact. On the physical side we can cite a list of possible health hazards that seem to occur more frequently in substandard housing, and in the case of a few of these hazards we have shown correlations with ill health as well as some causal mechanisms.¹ On the mental health side we have some rather vague ideas concerning the role of substandard housing in relation to the individual's self-image.² Only in the area of density and its effects do we seem to be making some progress, largely through work in the field of proxemics.³

¹Alvin Schorr, Slums and Social Insecurity (Washington: U.S. Government Printing Office, 1963), p. 14.

²Ibid., pp. 8-12.

³See for example, Edward T. Hall, The Hidden Dimension (Garden City, New York: 1966).

Density in itself, however, has little to do with the state of physical repair of the dwelling unit. Evidence which relates to physical condition per se to health and safety is relatively scarce and inconclusive.

On the other hand, if certain aspects of housing can legitimately be shown to be a public health problem, then health authorities should be mounting constructive programs to deal with this problem. Thus there is great need for additional work in the health field as it relates to housing. Such work can serve two purposes. First, we may be able to clarify the vague situation which exists at present, thus providing less room for the uncritical acceptance of traditional dogma. Secondly, we may be able to devise more positive programs to deal with the health problems which do in fact exist.

This thesis represents an attempt to attack one small segment of the problem by analyzing a set of home accident fatality statistics in relation to census data on housing quality. Home accidents are a logical subject to consider when attempting to relate housing to health, given the current lack of knowledge. As will be seen, home accidents are a public health problem of considerable magnitude. As such, home accident statistics deserve to be given much more systematic analysis than has been the case heretofore, with a view toward the development of meaningful preventive programs. Also, accidents are really about the only type of health problem related to housing for which statistical data is readily available, although even here the data is certainly not as suitable for analysis as it might be. (A discussion of some of its shortcomings will be found further on in the report.) Thus, given the current state of the art, time spent on analysis of the home

accident problem might be more fruitful than time spent on some of the more esoteric health problems that have been attributed to substandard housing.

Accidents -- The Problem

A brief consideration of general accident statistics quickly substantiates the idea that accidents of all types are a major public health problem. Accidents claimed approximately 105,000 lives in the United States in 1964.⁴ They are the fourth leading cause of death for the population as a whole (53 deaths per 100,000 people), ranking behind only heart disease (375), cancer (151), and vascular lesions (107).⁵ For all age groups between the ages of one and 37, accidents are the leading cause of death.⁶

The total cost of accidents occurring in the United States in 1964 was estimated by the National Safety Council to be \$16.7 billion.⁷ Of this figure, as shown in Table 1, the monetary cost of injuries from these accidents was estimated to be \$9.9 billion, or almost 60 percent of the total. The remainder is property damage. These figures represent a minimum estimate of the cost of accidental injuries, however. They include only the lost wages, medical expenses, and the overhead

⁴National Safety Council, Accident Facts (Chicago: Annual), 1965, edition, p. 3.

⁵Ibid., p. 8.

⁶Ibid.

⁷Ibid., p. 5.

Table 1
 Certain Costs of Accidental Injuries, 1964⁸
 (Millions of Dollars)

Cost	Moto- Vehicle	Work	Home	Public	Total
Wage Loss	2,200	1,350	800	800	5,000
Medical Expense	500	600	400	200	1,650
Overhead of Ins.	<u>2,600</u>	<u>650</u>	<u>10</u>	<u>10</u>	<u>3,250</u>
Total	5,300	2,600	1,200	1,000	9,900

cost of insurance, and not the unmeasurable costs of human suffering, discomfort, and inconvenience. Further, the data from which these estimates are derived includes only those accidental injuries serious enough to involve medical treatment and/or a restriction of activity for at least one day on the part of the person injured. The total number of persons actually injured far exceeds the number whose injuries involve measurable monetary costs.

Tables 2 and 3 give an idea of the total number of injuries of various types occurring annually which either require medical care or involve restriction of activity or both. Note that the left hand columns of the two tables correspond approximately. They do not agree exactly because of the differences in the definitions used by the National Safety Council and the National Health survey. Table 3, however, is essentially an expansion of the "bed disabling" category of Table 2 into a set of subcategories.

Table 2
 Estimated Number of Bed Disabling and Less Serious Injuries,
 Millions, Annual Average, July 1961 - June 1964,
 From National Health Survey Data⁹

Class	Bed Disabling	Not Bed Disabling		Total Injured
		Act. Restriction	Not	
All Classes	10.89	17.36	22.53	50.76
Motor Vehicle	1.29	0.95	1.07	3.31
Work	1.72	3.18	4.11	9.02
Home	4.26	7.33	10.54	22.13
Other	3.96	6.27	7.18	17.14

Table 3
 Accidental Injuries by Severity of Injury,
 Millions, 1964. National Safety Council Data.¹⁰

Class	Total	Temporary Disability	Permanent Impairment	Deaths
All Classes	10.30	9.85	0.37	0.105
Motor Vehicle	1.75	1.60	0.14	0.048
Work	2.05	1.95	0.08	0.014
Home	4.35	4.20	0.11	0.029
Public	2.25	2.20	0.05	0.018

Several important facts can be deduced from these tables. From the National Health Survey data we see that approximately 50 million persons suffer some form of accidental injury severe enough to restrict

⁹Accident Facts, 1965 edition, p. 2.

¹⁰Accident Facts, 1965 edition, p. 5.

activity or require medical attention each year. This amounts to roughly one-fourth of the total population. Further, sources within the National Health Survey indicate that only about half of the injuries actually reported to the survey are severe enough to be indicated in its tabulations.¹¹ Thus it appears that approximately one-half of the population of the United States suffers some form of accidental injury, however minor, each year.

Turning specifically to home accidents, we see that they do in fact constitute a large part of the overall problem. Although most of the headlines tend to be reserved for motor vehicle fatalities, it is evident that the home accident fatality problem is almost of the same magnitude. Home accidents claimed 28,500 lives in 1964, as compared with about 48,000 motor vehicle fatalities.

As seen from Table 1, home accident injuries rank relatively low in terms of measurable costs. When we turn to the minor forms of injury, however, home accidents begin to emerge as a major problem. Although home accidents rank below motor vehicle accidents in terms of deaths, permanent impairments, and measurable costs, they rank well above all other types, including motor vehicle accidents, in terms of temporary disabilities and total injuries. As seen from Tables 2 and 3, home accidents are responsible for approximately 43 percent of temporary disabilities and 45 percent of all accidental injuries. Home accident injuries of all kinds total more than twice the number of work injuries and almost seven times the number of motor vehicle accident injuries.

¹¹Philip S. Lawrence, "Collection of Data on Accidental Injuries," Public Health Reports (Volume 74, March 1959), p. 197.

Thus, if we may generalize about home accident injuries we may say that they are less serious than other types (such as motor vehicle accident injuries) but that they are far more prevalent. And although home accident injuries are in general less serious than other types, they nevertheless account for a substantial proportion of the fatalities and permanent impairments. Therefore home accidents appear to be a serious public health problem. This gives rise to the question of whether we can begin to develop a model of the home accident process and from this perhaps begin to find some solutions.

Accidents -- The General Model

Although the term "accident" tends to imply a lack of causal mechanisms, public health experts have come to view the problem as one in which there are, in fact, causes, although for any given accident the causes are multiple and the accident mechanism probabilistic rather than deterministic. The seminal work which seems to have caused a reorientation toward the view of a probabilistic model with multiple factors is one written by Professor John E. Gordon of Harvard in 1949. In the introduction to his article Gordon said:

If home accidents are primarily a public health problem, then that problem is reasonably to be approached in the manner and through the technics [sic] that have proved useful for the other mass disease problems. This includes first an epidemiologic analysis of the particular situation, an establishment of causes, the development of specific preventive measures directed toward those causes, and finally a periodic evaluation of accomplishment from the program instituted.¹²

¹²John E. Gordon, "The Epidemiology of Accidents," American Journal of Public Health (Volume 39, April 1949), p. 504.

Gordon saw the possibility of three fundamental types of factors in causation: those of the host (man); those of the agent (unprotected heights, exposed electrical wiring, faulty pavement, unguarded poison, etc.); and those of the environment (physical, biologic, socio-economic).¹³ Under this classification system the distinction between the physical environment and the agent does not seem entirely clear, but Gordon seems to mean by "physical environment" primarily such things as climate, weather, season, topography, terrain, etc., thus reserving the actual physical agent for a separate category. This discussion will make use of the system put forth by Gordon, with the exception that a category entitled "hazard" will be substituted for "agent." This category will include as subcategories both the agent itself and the mechanism which allows it to function in a harmful way, i.e., the hazard is seen as consisting of two parts, the agent and the mechanism.

It also might be worthwhile to think in terms of some sort of density function of hazards in the environment. This will be done later in the chapter.

Although no one has yet attempted to build a complete model along the lines of such a classification, conceivably we might some day be able to build a model of home accident incidence (or general accident incidence) given the appropriate data on individual characteristics, characteristics of the socio-economic environment, characteristics of the physical environment (in the sense of general environmental forces), and the geographic density of various types of environmental hazards.

¹³Ibid., pp. 509-512.

This will require a great deal more research than is presently being done. At present we have only some very general schematic conceptions of such a model,¹⁴ and the knowledge to begin to be more specific about a few limited aspects of it. For the most part these schematic conceptions do little more than indicate that the individual may have a higher or lower initial probability of having an accident, depending on his individual characteristics, that this probability may be modified as a result of the action of a multiplicity of environmental forces, and that the environment may contain a varying number of potentially hazardous situations, situations which might be thought of as presenting a better than average opportunity for actually having an accident given that the individual is at some level of propensity of "proneness."

Human Failure -- Individual Characteristics

A report by Operations Research Inc. lists the following as causes of human failure which may contribute to accidents:¹⁵

- (1) Risk taking behavior.
- (2) Inattention, distraction.
- (3) Visual, auditory, or other sensory deficiency.
- (4) Sudden incapacitation resulting from heart disease, etc., or severe pain.

¹⁴For example, schema presented by M. S. Schulzinger, The Accident Syndrome (Springfield, Ill.: 1956), p. 2, and Operations Research Inc., Analysis of Responsibility and Capability of the Public Health Service in Accident Prevention (Washington: U.S. Department of Health, Education, and Welfare, Public Health Service, 1961), p. 9.

¹⁵Operations Research Inc., op. cit., pp. 43-44.

- (5) Lack of adequate knowledge or experience to:
 - (a) recognize danger.
 - (b) avoid danger following detection.
- (6) Inadequate motor performance.
 - (a) physiological.
 - (b) skill.

If we examine such a list, it is evident that certain of these causal factors will be more prevalent among some population groups than among others. Sensory deficiencies and inadequate motor performance, for example, are primarily a problem among the elderly.¹⁶ Lack of adequate knowledge or experience to recognize and avoid danger, on the other hand, is a problem confined largely to the very young. Risk taking behavior seems to vary with age and sex, as well as with other individual characteristics. It is also affected by a multitude of environmental forces operating through the individual psyche. An example given by Malfetti will illustrate this idea. Although the example concerns driver behavior, an analogy could be constructed for any other environmental and behavioral situation.

Malfetti views driver behavior as being at any moment a point on a continuum, with safe driving at one end and unsafe driving at the

¹⁶Gordon J. Azar and Alfred H. Lawton, "Consequences of Physical and Physiological Change with Age in the Patterns of Living and Housing for the Middle-Aged and Aged," Patterns of Living and Housing of Middle-Aged and Older People (Washington: U.S. Department of Health, Education, and Welfare, Public Health Service, 1965), pp. 19-26. Also Gordon J. Azar and Alfred H. Lawton, "Sensory and Perceptual Changes that May Influence Housing Needs of the Aging," Ibid., pp. 11-15.

other. He sees the initial position and stability of this point as being determined largely by the characteristics of the individual.¹⁷

He then goes on to say that:

In constant interaction with these characteristics are hundreds of forces, usually transient, which relate to a particular driving episode. Examples are the behavior of other drivers, presence of enforcement officials, signals, flow of traffic, and conditions and characteristics of both highway and vehicle. Also influencing the driver are such factors as where he is going, how much time he has to get there, what he has been eating and drinking, how he is feeling, what he is thinking about, looking at, and listening to, and his immediate relationship with his passengers, his wife, his boss, and other persons important in his life.

As these forces interact with one another and with characteristics of the individual, they may move the point representing driver behavior toward the safe or the unsafe end of the continuum.¹⁸

A great deal of the accident literature has focused on the question of "accident proneness." The question of accident proneness has arisen because of the incidence of accidents has been found to be closely related to personality or psychological factors. A resume of literature on this question can be found in either Malfetti (cited above) or Schulzinger.¹⁹ Both are in agreement that psychological factors are of prime importance in the causation of accidents. Both also agree, however, that inherent accident proneness -- in the sense of one person being accident prone while others are not -- does not exist.

¹⁷James L. Malfetti, "Attitudes and Safety in Recreation," Public Health Reports (Volume 78, June 1963), p. 478.

¹⁸Ibid., p. 479.

¹⁹M. S. Schulzinger, The Accident Syndrome (Springfield, Ill.: 1956).

Rather, the mental maladjustment which seems to lie at the base of risk taking behavior is usually fleeting or only mildly prolonged and it occurs in all segments of the population in response to the interaction of the individual's personality with a variety of transient environmental forces.

An interesting question arises, however, as to the variation in levels of risk taking behavior with the average levels of various physical and socio-economic forces in the environment over the long term. We might hypothesize, for example, that the level of risk taking varies with the degree of irritability caused by the level of physical stresses such as noise, or with excessive heat or cold. Another variable which might affect risk taking behavior through the mechanism of increased irritability is residential density. Finally, risk taking might vary with general socio-economic level. If so, any analysis which attempts to relate home accidents to housing quality will be biased unless the variation in risk-taking with socio-economic level is taken into account.

None of the above hypotheses have been tested, but the possibility remains that the level of risk taking may vary in important ways with the levels of various environmental variables just as it seems to vary with age and sex.

One study of which the author is aware which begins to impinge on this type of question is concerned with expressive self-testing in driving.²⁰ The need to test one's self in this fashion would seem to

²⁰John M. Roberts, Wayne E. Thompson, and Brian Sutton-Smith, "Expressive Self-Testing in Driving," *Human Organizations* (Volume 25, Spring 1966), pp. 54-63. Also Cornell University Center for Housing and Environmental Studies, Division of Urban Studies, Article Reprint No. 17.

be one among many emotional states which might lead to a higher level of risk taking behavior. In the particular study cited, the authors attempt to relate high self-testing to a variety of other attitudes as well as to class membership. Although the types of attitudes with which their survey was concerned are not particularly relevant to the subject being considered in this thesis, surveys of this type might be a means of relating psychological factors thought to be important in accident causation to various environmental variables of the types mentioned above.

Environmental Hazards

We know that a person, if for some psychological reason his propensity to have an accident is high enough, can have an accident even in the relative safety of a padded cell. Nevertheless, for the overwhelming majority of people who have accidents the propensity to have them is not that high. Thus we might hypothesize that a person who is emotionally very upset on a given occasion might be liable to have an accident almost anywhere, while a person with a lower propensity might require a very dangerous situation in order to have an accident. If we are willing to accept this strictly qualitative idea, then the environmental situation which is considerably more dangerous than that considered "normal" might be thought of as a hazard. Although accidental injuries can occur anywhere, our hypothesis in essence says that, all other things being equal, they are more likely to occur where the number of hazards is large.

It is appropriate to divide the hazard into two elements, because data which give us only the agent of injury without also giving us the physical mechanism of the accident is of little value when we attempt to devise preventive programs. One or two examples might be appropriate to illustrate the distinction between the two. Suppose we have an accident fatality in which a fall is considered to be the agent of injury. This information is relatively useless unless we are also given the mechanism of the accident, which might be something like "slipped on loose rug," or "tripped on defective stair," or "weak railing broke, allowing person to fall." For another example, suppose we have the accidental poisoning of a young child. The agent might be considered to be the poison itself. The mechanism, however, might be either "was left in the open after use by parent," or "storage place was accessible to child." These are two entirely different situations, one involving the layout of the home itself and the other involving neglect on the part of the parent.

The distinction between mechanism and agent used here seems especially meaningful because all housing can be expected to have a variety of agents, while not all housing has the type of situations which would correspond to the mechanism being a fault of the house when an accident occurs.

Although all housing can be assumed to contain hazards, it seems reasonable to expect that so called "dilapidated" housing has in general a greater number of such hazards per dwelling unit. Another way of expressing this idea is to say that we expect dilapidated housing to have a higher density of hazards. If this is so, there may

thus be more opportunities for the individual to have accidents if he is at a given level of accident propensity.

Although it has not been shown that either (1) a higher density of hazards actually prevails in dilapidated housing or (2) the greater number of opportunities actually results in more accidents and injuries, this assumption seems to be one of the underlying rationales for our concern with substandard housing.

Conclusion

So far we have seen that home accidents do in fact constitute a serious public health problem, and that we have only a very general accident model. This model is very hypothetical and has not yet given us the insights which we need if the problem is to be attacked. Thus we need much more research in order to begin to fill in the gaps and answer some of the questions which the model poses.

The model has been presented here for two purposes. The first is simply for the general information of the reader. The second is so that we may make use of it in order to develop some hypotheses about the meaning of the rest of the thesis.

The remainder of this thesis will be primarily concerned with the description of an attempted correlation of home accident fatality rates with levels of housing dilapidation in a set of selected counties. Thus the generalized model set forth here should be kept in mind so that we may eventually return to it with the idea of constructing a set of possible hypotheses about what such a correlation might mean.

CHAPTER II
THE DATA AND THE ORIGINAL REGRESSION

Introduction

Chapters II and III describe an attempted correlation of home accident fatality rates with housing quality data by counties for the State of Kansas during the period from 1950 to 1957. The hypothesis is that accident fatality rates are correlated with housing quality.

The idea behind the analysis is that before an assertion is made that dilapidated housing is unsafe and mount programs to eliminate it, we should attempt to test this assertion. This involves first testing the idea that a correlation exists at all, and either accepting or rejecting this hypothesis at some level of statistical significance. If this hypothesis were tested with several different sets of data and repeatedly rejected, then that would be the end of the matter. If, on the other hand, the hypothesis was repeatedly accepted we should then return to the home accident model and begin to formulate and test hypotheses about why the correlation exists. None of this has yet

been done, and thus this thesis attempts to deal with one small segment of the process.

There was no particular reason for choosing the set of data used here except that it was conveniently available and as applicable to the hypothesis under consideration as any other available data. This does not mean that the data was entirely satisfactory, and in particular one of its limitations should be recognized and kept in mind when interpreting the results. That limitation is that the data concerns home accident fatalities only, and not home accident injuries in general. We have no proof that fatality data can be used as a surrogate for data on injuries of all types, and thus it should not be used as such. Unfortunately the fatality data is the only data available because the regular collection of data on other types of injuries by geographical subdivisions is not carried out by any governmental agency.

The Fatality Data

The fatality data used in the analysis was taken from the Kansas Accidental Death Report, a biennial volume of accidental death data for the State of Kansas.²¹ Four volumes were used, giving data for the eight consecutive years from 1950 to 1957. Data is tabulated in a variety of ways in the reports: by counties; by age groups; and by types of agent; etc. Since data was available by counties, it was decided to attempt to correlate this data in some way with a measure of housing quality by counties.

²¹Kansas State Board of Health, Division of Vital Statistics, Kansas Accidental Death Report (Kansas State Board of Health: 1952, 1954, 1956, 1958).

The Dilapidation Data

Data on the percentage of "dilapidated" housing was taken from the U. S. Census of Housing. According to the census definition, a dwelling unit is to be considered dilapidated if it meets any one of the following criteria:²²

- (1) Having one or more critical deficiencies, such as: holes, open cracks, rotted loose, or missing materials over a considerable area; substantial sagging of floors, walls, or roof; extensive damage by storm, flood, or fire.
- (2) Having a combination of minor deficiencies present in sufficient number and extent to give evidence that the unit does not provide adequate shelter or protection from the elements or is physically unsafe, such as: holes, open cracks, rotted, loose, or missing materials over small areas; shaky or unsafe porch, steps, or railings; damaged chimney; broken, loose, or missing stair treads, balusters, or railings; deep wear on doorsills, doorframes, steps, or floors.
- (3) Having inadequate original construction, such as make-shift walls, lack of foundation, and dirt floors.

Data was given in the census tables on total number of dwelling units occupied, the total number of dilapidated units, and the number of vacant dilapidated units. It was felt that only occupied units would be relevant, since quite obviously no accidents would have occurred in empty units. Therefore the total number of occupied dilapidated units was calculated. This was done for each county for the years 1950 and 1960.

²²U.S. Bureau of the Census, U.S. Census of Housing: 1950. Volume I, General Characteristics, Chapter 16, Kansas (Washington: U.S. Government Printing Office, 1952), p. ix.

Since the fatality data covered an eight year time span and the percentage of dilapidated units might alter considerably within that time, it was felt that a statistic which would at least approximate an average percentage of dilapidated units for the period was needed. There was, of course, no way of knowing what the true average actually was, so it was assumed that both the number of occupied units and the number of occupied dilapidated units varied linearly during the period from 1950 to 1960. The statistic actually used was the percentage of dilapidated housing in 1955. Given the above assumptions, this was calculated as $(OD_{1950} + OD_{1960}) / (OU_{1950} + OU_{1960})$, where OU equals the number of occupied units in the particular year and OD equals the number of occupied dilapidated units. The decision to use 1955 was made because, although 1955 was not the exact midpoint of the period under consideration, it was close enough to the midpoint so that the slight variation in relative percentages of dilapidation that might obtain would not affect the final result significantly and because the choice simplified the necessary calculations.

One question which arose was that of the reliability of the census data on dilapidated housing. Although this question cannot be answered here, one wonders about variations in perception among census takers. The question came to mind because of a very interesting discrepancy between the housing census data of 1950 and that of 1960. As is well known, an additional classification called "deteriorating" was added to the census of housing in 1960. In the case of the Kansas data being used here, this addition seems to have had the effect of reducing the number of units considered "dilapidated" to a level far

that which might have been expected on the basis of the 1950 levels. It is clear that many of the units considered "dilapidated" in 1950 were considered to be "deteriorating" in 1960. Thus, it is apparent that a large number of borderline cases exists where the judgment of individual census takers might vary markedly.

It is also evident that a whole range of housing quality exists and that the distribution of dwelling units throughout this range may or may not be reflected by the distribution between the two categories "dilapidated" and "not dilapidated." It might be worthwhile to attempt a study on how well the number of dilapidated units reflects the average level of housing quality in a set of geographic units. In the present case it is tacitly assumed that the number of dilapidated units is a reasonable indicator of the general level of housing quality.

With regard to the question raised previously, it is to be expected that when a large number of observations are used (105 counties in this case) the variation in judgment of census takers will tend to cancel out. The same can be said when larger units are used, because larger numbers of census takers are required for these units.

Populations and Death Rates

Data on the populations of the various counties was obtained from the U. S. Census of Population, 1950 and 1960. In general there were relatively large population shifts within the state during the period in question, with most of the small rural counties showing declines and the larger, more highly urbanized counties showing substantial gains. Again, for lack of better information, it was assumed that

each county shifted population in a linear fashion and on that basis a population was computed for the midpoint of the period in question. Raw death rates for the counties were computed by dividing these populations into the accident totals for counties which had been tabulated from the Kansas Accidental Death Report. Accidental deaths by counties can be found in Appendix Table I and the raw death rates, along with the populations and numbers of deaths used in computing them, can be found in Appendix Table II.

Home Accidents and the Binomial Model

The first question which arises when attempting to analyze a set of data is that of the appropriate analytical methods to use. In this case the resolution of this question was aided considerably by viewing the home accident mechanism as acting in a manner similar to that of a bernoulli process and yielding home accident death rates for counties in a manner similar to the binomial model.

The binomial model is applicable to situations with respect to sums of random variables having only two possible values or categories. In this case the resident of a given county during a certain time period may either (1) have died as the result of a home accident, or (2) not have died as the result of a home accident. Obviously each of these categories includes a number of subcategories. The first can be broken down by type of accident (mechanism and agent). The second includes "die, but not of a home accident;" "have a home accident, but not die;" and "not have a home accident and not die" -- but this does not matter. The situation can be described by a random variable having

only two possible alternatives: to die as the result of a home accident or not.

The second condition defining the binomial model, that for any single trial -- or for any single individual in the population during the given time period -- the probability of dying as a result of a home accident (the bernoulli parameter, p) is the same, presents a more difficult problem. This condition requires that each individual in the population -- in our case each individual residing within a given county during a certain period of time -- has the same probability as each other individual in the county of being the victim of a home accident death. As seen in Chapter I, this is not the case. We know that the probability varies with certain individual characteristics such as age; that this probability may be altered by a myriad of environmental forces; and that even if a set of individual characteristics and general environmental forces is specified we can talk only of the average probability for individuals within that homogeneous group.

Thus we must recognize that the second condition is not satisfied. Nevertheless, it seems reasonable to make the assumption that the result obtained from a process in which the individual observations are subject to different probability parameters is similar to the result which would be obtained if each individual observation were subject to the average probability parameter of the group. Thus we are assuming that if we had a large group of m counties, all of population size n and all subject to the same average probability parameter (on the part of the individual person), the resulting distribution of home accident deaths per county (and consequently home accident

death rates) would be a binomial probability distribution resulting from a binomial process with parameters p and n .

As an example, suppose that there existed a hypothetical set of counties each with a population of n people. Further, suppose that in each of these counties during a given time period each individual had a probability p of suffering a home accident death. For this hypothetical situation the binomial probability of exactly u home accident deaths occurring in a given county, where u can assume any of the values $0, 1, 2, \dots, n$, would be given by the expression

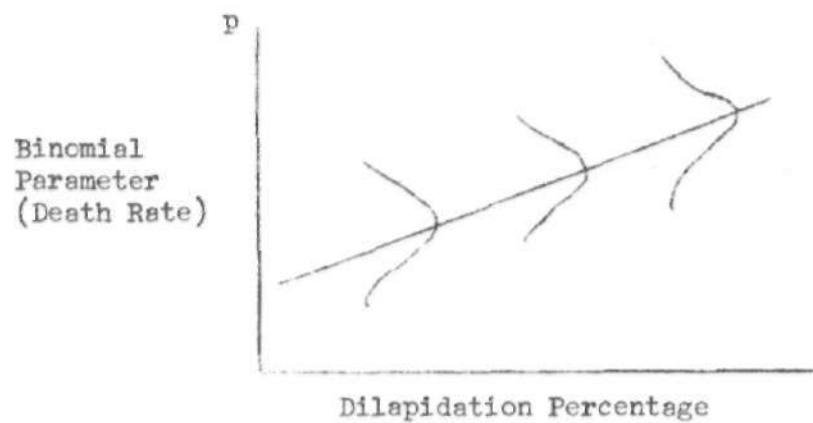
$$\Pr(u) = \frac{n!}{u!(n-u)!} p^u q^{n-u}$$

where $q = 1-p$. Substitution of the appropriate values into this expression would yield the binomial probability of exactly u home accident deaths occurring, convertible into death rates, for the hypothetical set of counties of a given population size.

As a test of the idea that home accident fatality rates follow the binomial model, binomial distributions were computed for two theoretical sets of counties, one with $n = 2000$ and $p = .0016$ and one with $n = 10,000$ and $p = .0016$. These were compared with a set of 41 counties with an average $p = .0016$ drawn from the Kansas data. These are shown in Appendix Tables IV through VI the real life distribution, after some arbitrary smoothing, is remarkably similar to the theoretical distributions. Thus it would seem that the distribution of deaths and death rates among counties does follow a process similar to the binomial model. This idea is important because it yielded some insights which were important later in the analysis.

Choice of Analytical Techniques

The hypothesis that a correlation exists between housing quality (as measured by the percentage of dilapidated housing) and home accident fatality rates, when put in terms of the binomial model, says that the binomial parameter p should vary as a function of the percentage of dilapidated housing. In the situation at hand in this thesis, we have 105 counties, each with its own percentage of dilapidated housing and its own hypothesized binomial parameter p . If for each of these counties we had a group of others identical to it, the death rates within each group of counties should be binomially distributed according to the appropriate parameters of that group, p_i and n_i . If we took the data for these 105 counties and ran a linear regression (assuming that p varies linearly as a function of dilapidation percentage), the regression equation should pass through the expected means of the 105 hypothesized binomial distributions. This idea is shown in the accompanying diagram. The diagram shows three of a potentially infinite number of binomial distributions which could exist, or of the 105 which are hypothesized to exist in this case.



Given the nature of the hypothesis as shown above, there are two possible methods of testing it. One method would be to compute the average home accident death rate, for the entire state and attempt to determine whether the range of home accident death rates by county which actually existed would be possible using the statewide average rate as a binomial parameter. This could conceivably be tested at some level of significance. There would be several problems involved, however, which would make such an analysis very difficult. There would be a problem in choosing an appropriate county size parameter, n , since the actual county sizes vary over a wide range. Secondly, the home accident process does not meet the conditions of the binomial model so exactly that we could have any confidence in the result of such an analysis. Third, the analysis would not yield a result which would tell us something about the actual relationship which might exist between the two variables, i.e., an equation. Thus this method was rejected as not being suitable for the situation at hand.

The method actually chosen for use in the analysis was the method of least squares linear regression. Although there proved to be problems with this method also, discussion of these will be deferred until later in the chapter.

The actual analysis contained the following sequence of steps. Initially, a linear regression was run on the raw data which was presented earlier in the chapter. As this was being done, some further thinking about the problem led to the conclusion that a regression on the raw data would not take certain factors into account and thus would not be entirely valid. The analysis then continued through a

process of determining the appropriate manipulations to make on the data in order to make the regression more valid, making these manipulations, and running regressions on the adjusted data in order to obtain a more valid result. This sequence will be described through the remainder of this chapter and Chapter III.

The Initial Regression

As mentioned previously, the total number of deaths by counties, dilapidation percentages, and raw home accident death rates are shown in Appendix Tables I, II, and III. The home accident death rates were plotted against dilapidation percentages, as shown in Appendix Figure I, and a straight line was fitted to these points by means of a least squares linear regression. The regression equation was found to be $y = 16.12 + 0.174x$, thus indicating a slight correlation between substandard housing and home accident fatality rates. The units used in the equation are deaths per 10,000 population per eight years (y) and percentage of dilapidated housing (x), respectively.

In spite of the slightly positive slope of the regression equation, the standard error of estimate was found to be 6.27 and the coefficient of determination (r^2) was found to be only .00365, indicating that only a little more than one-third of one percent of the variation was explained. A t-test showed that the level of significance was only .55, indicating that our result could easily have occurred with absolutely no correlation between the two variables 55 percent of the time. Thus the outlook for further analysis of the data was not particularly favorable.

For several reasons, however, the results of the initial regression were not considered to be conclusive and further analysis was done. These reasons included the following.

Problems with Use of Regression on Raw Data

There are two major reasons why the use of a regression on data concerned with raw home accident death rates by counties is not valid. These are: (1) the fact that the age distribution of the population varies markedly among counties; and (2) the fact that counties are not all of the same population size. The remainder of this chapter will consist of a discussion of the problems which these two facts pose.

Variation in County Population Sizes

The distribution of population sizes in the 105 Kansas counties is shown in Table 4.

Table 4

Population Size	No. Counties
0 - 5,000	20
5,000 - 10,000	33
10,000 - 15,000	18
15,000 - 20,000	10
20,000 - 25,000	8
25,000 - 50,000	11
50,000 - 100,000	2
Over 100,000	3

As can be seen from this table, the majority of counties are clustered at the lower end of the population size scale, with approximately half of them containing less than 10,000 inhabitants each. There are only a very few large counties. In the simple regression the counties are treated equally, with the smallest county receiving as much weight in the analysis as the largest county. There are reasons, with respect to both of the variables being used in the analysis (dilapidation percentages and home accident death rates), for considering this equal weighting to be invalid.

With respect to the dilapidation percentages, it will suffice to quote the following from the U.S. Census:²³

There is evidence that some enumerators obtained too large or too small a count of dilapidated units. Consequently, data on condition for small areas, which depend on the work of one or two enumerators, are subject to a wider margin of error than is to be expected for larger areas, where the data represent the returns made by several enumerators.

With respect to the home accident death rates, Table 5 shows standard deviations relative to expected death rates for the theoretical binomial distributions for sets of counties of various population sizes and $p = .0016$. The two right hand columns are shown in terms of deaths per 10,000 population per eight years in order to correspond with the data being used in our analysis. The figures in these two columns indicate that when large population units are used the actual death rate is likely to fall much closer to the expected death rate.

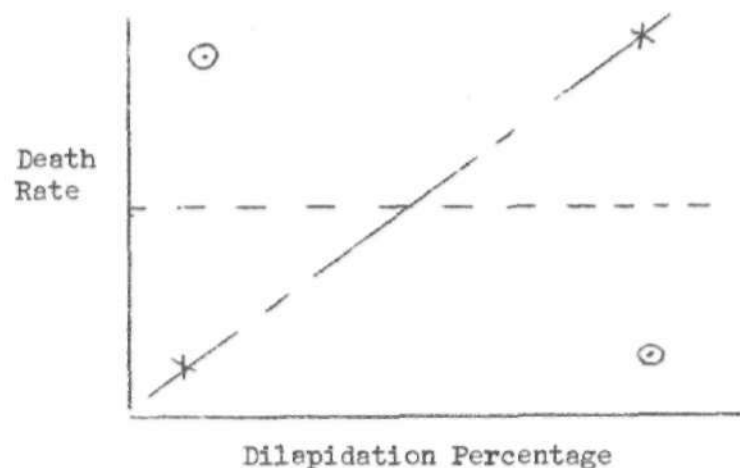
²³U.S. Bureau of the Census, U.S. Census of Housing: 1950. Volume I, General Characteristics, Chapter 16, Kansas (Washington: U.S. Government Printing Office, 1952), p. ix.

Table 5
Standard Deviations of Binomials, $p = .0016$

Assumed County Size	Expected Deaths	Standard Deviation	Per 10,000 per Eight Years	
			Expected Deaths	Standard Deviation
3,000	4.8	2.2	16	7.3
5,000	8.0	2.8	16	5.7
10,000	16.0	4.0	16	4.0
50,000	80.0	8.9	16	1.8
100,000	160.0	12.6	16	1.3

Therefore, the death rate for a large county can be expected to yield more information than is yielded by the death rate for a small county.

There is another way of looking at this question that will be illustrated with a simple example. Suppose that in the accompanying diagram the x's represent counties of one million population each and the circled dots represent counties of one thousand population each. If we wished to determine the relationship between the death rate and percentage dilapidation, a simple regression treating the counties equally would give a horizontal line, i.e., no relation. This result is obviously absurd. Suppose that we could break each of the large counties down into a thousand units of one thousand population each -- then the mean death rate and mean percentage dilapidation for each of these groups would fall in exactly the spots already indicated by x's, and the regression line through the total 2,002 points would be a diagonal as shown on the diagram. This result assumes, of course, that the smaller units would behave in the manner assumed by the binomial model.



Thus there is certainly a question as to whether counties of varying sizes should be treated equally as in a simple regression, and something can be said for giving the larger counties a weight more in line with the information which they contribute. This might mean either using only the larger counties in the analysis or aggregating the counties into larger population units in order to give each county a weight commensurate with its population size.

Variations of Population Age Distribution

The second problem to be considered is the variation of the age distribution of the population among counties. This was not taken into account in the initial regression and thus was one of its major weaknesses. Table 6 shows age group-specific home accident death rates derived from the Kansas data, in deaths per 10,000 population per eight years. It is immediately evident that the death rates of counties should vary widely according to the age compositions of their population.

Table 6
Age Group -- Specific Death Rates

Age Group	Deaths/10,000/8 years
0 - 44	20.5
5 - 44	3.68
45 - 64	8.35
65 and over	97.1

A home accident death rate of 97.1 would not be at all out of line for a county whose population was virtually all over 65 years of age. Thus any analysis which compares death rates by counties must take the age factor into account.

Conclusion

This chapter has described the data being used and the initial linear regression done on that data, as well as some of the problems inherent in both. The following chapter will describe the adjustments which were made in an attempt to cope with these problems.

CHAPTER III
FURTHER MODIFICATIONS

Introduction

As seen in the previous chapter the initial attempt to relate housing quality with home accident fatalities, a simple regression performed on the raw data, was not an outstanding success. Although the regression showed a slight correlation, the regression equation was found to explain virtually none of the variation in the data and the standard error of estimate was quite large. Aside from this, there also remained a question as to whether a regression was appropriate at all given the wide variation in population sizes of counties. Finally it was felt that the variation in age composition among counties should also be taken into account.

In order to correct these deficiencies some further steps, to be described in this chapter, were taken. Chief among these was the use of an "equivalent population" concept to adjust the population sizes of counties in order to remove the effect of variations in age composition.

Also used was the idea of aggregating the counties into larger groups, both to cut down on the amount of variation in the data and to remove the effect of variations in county size.

The Equivalent Population

Two means of adjusting for variations in age composition, essentially two sides of the same coin, were examined. These can be characterized as the concept of "expected deaths" and the concept of "equivalent population," respectively.

As will be remembered from Chapter II, the home accident death rate varies markedly among age groups in the population. The concept of expected death begins with the assumption that these age group specific death rates, computed on the basis of the statewide population, should apply equally well to the individual county populations if all other factors are held constant. Thus if we apply the statewide age group specific death rates to the age group cohorts of an individual county and sum the number of deaths obtained, we get an "expected" number of deaths for the county, i.e., the number of deaths to be expected in that county if the statewide age group specific death rates remained constant, on a geographic basis, over the entire state. An example of this calculation is shown.

Age Group	Pop. (10,000's)	Statewide Rate	Exp. Deaths
0 - 4	6.64	20.5	136
5 - 44	34.14	3.68	126
45 - 64	10.17	8.35	85
65+	<u>4.48</u>	<u>97.1</u>	<u>435</u>
Total			782

The number of expected deaths can then be compared with the number of deaths which actually occurred, i.e., the ratio of actual to expected deaths can be computed. The example shown is the actual computation for a group of Kansas counties which had been aggregated. In this case the actual number of deaths was 704 as compared to the 782 expected, giving a ratio of actual to expected of 0.90.

If this calculation were done for all counties in the state, the ratios could then be plotted against the corresponding percentage rates of dilapidation and a regression equation fitted to the data. The result, however, would obviously be in terms of the ratio of actual to expected deaths. While this ratio is conceptually useful, it was felt that in the current analysis it would be worthwhile to attempt to keep the results in terms of death rates. Therefore, another method was devised which did this while performing the same function, that of adjusting the data for variations in age composition. This procedure can be called the method of "equivalent population."

In order to perform the "equivalent population" computation for a county with a given age group distribution of population, the population of each age group cohort is multiplied by the ratio of the statewide home accident death rate for that cohort to the statewide home accident death rate of the population as a whole. This operation gives an "equivalent" population equal to the population which, at the average statewide death rate, would expect the same number of deaths as the original population at its cohort specific death rate. For example, since the population group over 65 has a home accident death rate of 97.1 while the population as a whole has a home accident death rate of

16.34, the over 65 cohort must be multiplied by the ratio $97.1/16.34 = 5.93$ in order to obtain its equivalent population.

When these products are summed they give an "equivalent" population for the county. Note that this "equivalent" population will differ in size from the original population of the county and, when applied to the actual number of deaths in that county, will inflate or deflate the average (non cohort specific) home accident death rate for that county according to whether the original age distribution of the county's population would tend to bias the actual number of deaths either upward or downward. For example, a county with an abnormally large segment of its population over 65 would tend to have a much greater number of deaths than would be expected if the average statewide death rate were applied to its population, simply because the elderly population experiences a much higher death rate statewide. In this case the equivalent population would be larger than the actual one, thus cancelling out the effect of differential age distribution. An example of this population calculation is shown, using the same aggregated group of counties as before.

Age Group	Pop. (10,000's)	Ratio	Eq. Pop
0 - 4	6.64	1.255	8.3
5 - 44	34.14	0.225	7.7
45 - 64	10.17	0.51	5.2
65+	<u>4.48</u>	<u>5.93</u>	<u>26.7</u>
Total	55.4		47.9

Note that in this example the equivalent population came out to be only 47.9, as opposed to the actual population of 55.4. The ratio of equivalent to actual was 0.865. To see the relationship to the expected deaths method, note that multiplication of the equivalent population by the statewide average death rate (47.9×16.34) gives the same number of expected deaths (782) as obtained before. Also, if we divide the actual number of deaths (704) by the equivalent population we get a death rate of 14.7, which when divided by the statewide average death rate of 16.3, gives us a ratio of 0.90, the same as that of actual to expected deaths. Thus the two methods are essentially two sides of the same coin.

Use of Equivalent Population with 16 Largest Counties

Since the equivalent population computation involves a considerable amount of time, the use of the concept on all 105 counties was judged to be unfeasible. A decision was made to use only those counties with over 25,000 population. It was felt that if this initial test showed that the population adjustments had significant effects, i.e., indicated that there were indeed large variations in the age group composition of counties, the entire set of 105 counties could then be aggregated into larger units and equivalent populations calculated for these.

The decision to use only counties with over 25,000 population resulted in a set of 16 counties containing approximately 56 percent of the state's total population. The data for these counties, both before and after the equivalent population adjustment, is shown in Appendix

Table VII. Regression equations were fitted to both sets of data by means of least squares linear regression.

The plotted sets of data, with regression lines, are shown in Appendix Figures II and III. The unadjusted data yielded the equation:

$$y = 6.15 + 2.29x,$$

with a standard error of estimate of 3.4 and a coefficient of determination (r^2) of 0.525. Through use of the t-test the correlation coefficient was found to be significant at the .002 level. The adjusted data yielded the equation:

$$y = 10.8 + 1.3x,$$

with a standard error of estimate of 1.9 and an r^2 of 0.55. The correlation coefficient was found to be significant at the .001 level.

While both correlation coefficients were found to be significant, it is interesting to note that the equations themselves were substantially different, as were the standard errors. Use of the equivalent population took a substantial amount of error out of the data.

Since the equivalent population adjustment did have such a substantial impact in this case and because data which has been adjusted in such a manner is obviously preferable to unadjusted data from a theoretical point of view, it was felt that the adjustment should be made in some fashion on all the data, even if doing the computations for all the individual counties was not feasible. Thus it was decided to aggregate the counties into groups on the basis of percentage rates of dilapidation and to perform the equivalent population adjustment on these groups.

Aggregating the Data

The data was aggregated in the following fashion. Four levels of dilapidation percentages (0-2.99; 3.00-4.99; 5.00-6.99; 7.00 and over) were chosen. The idea of aggregating on this basis was that it would give the aggregate groups the widest possible range of dilapidation percentages.

Once the counties had been grouped on this basis, the next step was to obtain age group populations for them. This could only be done by adding together the individual figures for each county to obtain totals for the entire group for both 1950 and 1960 and then interpolating on the assumption of a linear change during the period in question. A set of age group cohorts for each group of counties was obtained in this fashion.

Likewise with dilapidation rates, the individual rates for counties were not simply averaged. Rather, the individual numbers of occupied units and occupied dilapidated units were summed and the total occupied divided into the total occupied dilapidated in order to obtain a dilapidation percentage for the group as a whole. Since the county sizes varied so widely, to have simply averaged the individual rates would have been to give an invalid result. The dilapidation data for these aggregate groups is shown in Appendix Table VIII.

Death rates for these sets of counties were calculated on the basis of the original populations and on the basis of equivalent populations. The results, derived from data which can be found in complete form in Appendix Tables VIII and IX, are summarized in Table 7. Straight line equations were fitted to both of these sets of data (unadjusted and adjusted) by the method of least squares linear regression.

Table 7

Dilapidation Group	Aggregate % Dilapidation	Deaths	Pop.	Rate	Eq. Pop.	Adjusted Rate
0-2.99	2.21	704	55.4	12.7	47.9	14.7
3.00-4.99	3.85	1072	65.6	16.3	67.7	15.8
5.00-6.99	5.97	1136	60.4	18.8	63.9	17.8
7.00 and over	8.44	372	20.0	18.6	21.6	17.2

In the case of the unadjusted data the regression equation was found to be:

$$y = 12.0 + 0.90x,$$

with a standard error of estimate of 1.15 and a coefficient of determination (r^2) of 0.73. Use of the t-test showed the correlation coefficient (r) to be significant at the 0.16 level. The adjusted data yielded the equation:

$$y = 14.2 + 0.425x,$$

with a standard error of 0.66 and an r^2 of 0.70. The correlation coefficient was found to be significant at level 0.18.

This last equation should be considered to be the final result of this analysis. It has at least two advantages over the original regression equation of Chapter II in that (1) it takes into account the variations of age composition of the population among counties, and (2) it gives each county a weight more nearly commensurate with the information which that county has to offer, reflecting the fact that a very large county yields much more information in that it can be expected to fall much closer to the mean (for any given probability p) and thus should be given an additional weight in the analysis. This equation

should thus be a reasonable approximation of the result which would have been obtained if a regression had been run on the individual county data with (1) an equivalent population adjustment having been made on each individual county, and (2) the counties weighted according to population size.

The standard error and coefficient of determination obtained here, of course, apply only to the aggregate units used. If the individual county data had been used, even if adjusted and weighted in the fashion described above, the standard error would undoubtedly have been larger and the equation would not have explained as much of the variation. Both measures, however, would have been considerably better than those obtained in the original regression, for at least three reasons: (1) the regression line itself would have had a greater slope; (2) the data would have been weighted in favor of the larger counties, which tend to have less variation from the expected value; and (3) the equivalent population adjustment, as shown in both the regression on the 16 largest counties and the regression on the aggregated data, takes a considerable amount of error out of the data.

Our equation ($y = 14.2 + 0.425x$) says that the average equivalent home accident death rate (the mean or expected value of the binomial distribution for a set of counties with a given percentage of dilapidated housing and a population pyramid similar to that of the state as a whole) is equal to 14.2 plus 0.425 times the percentage of dilapidated housing. The units are deaths per 10,000 population per eight years. Assuming for the moment that we accept this equation with no reservations, it might be interesting to examine some of its implications.

Suppose, for example, that the state had no substandard housing -- presumably the home accident death rate would then have been 14.2 per 10,000 per eight years, as opposed to 16.3. This amounts to a drop of about 13 percent, or approximately 420 fewer home accident deaths (about 50 per year) over the eight year period.

Extending the argument somewhat further, from the regression equation we see that counties with no dilapidation should have an average death rate of 14.2 while counties with dilapidation rates of 100 percent should have an average death rate of 56.7. If we wish to extend the meaning of the analysis this far, these figures tell us that the person living in a dilapidated dwelling unit is four times as likely to suffer a home accident death as the person living in a non-dilapidated unit.

Extension of the meaning of our regression equation in this way is unwarranted, however, and thus should not be taken too seriously. Although a correlation of home accident death rates with percentages of dilapidation was found, the correlation coefficient was not found to be highly significant and it was obtained through the use of a limited set of data. Thus as far as the implication of the regression equation is concerned it would be more appropriate to say simply that it might be worthwhile to attempt to repeat the analysis on other sets of data. If correlations are found consistently with many sets of data, we should then begin to look for causal mechanisms.

Conclusion

This chapter completes the description of the analysis done. Although the analysis left many questions unanswered and was rather unsophisticated, it is felt that it was a significant improvement over the use of a simple regression with raw data and showed that, with the set of data being used (i.e., in one state during one period of time and under certain assumptions), there was some evidence of a correlation between home accident fatalities and dilapidated housing.

CHAPTER IV
CONCLUSION

The work on this thesis was undertaken with two main themes in mind. The first of these was to gain a general understanding of home accidents, the types of elements which make up the individual home accident sequence, and the nature of the home accident statistics which result when these individual events, spread out over time and space, are aggregated. The second was to take one small segment of the problem, i.e., one of the many unanswered questions, and begin to make progress toward answering that one particular question. In this case the question to be answered was whether or not a correlation exists between housing quality, as measured by census data on dilapidated dwelling units, and home accident fatalities. There was necessarily a great deal of interaction between the two themes, in the sense that, in order to make any meaningful statements about meanings of such a correlation or to be able to go beyond a simple regression analysis of the raw data, the process of working on the second theme required a simultaneous consideration of the first.

With regard to the first theme there are several important points which the reader should by now be aware of. Perhaps the most basic of these is that an "accident" is rarely an accident, at least in the commonly accepted meaning of that term. Most accidents are the result of a complex set of interacting factors and forces, both human and environmental. As seen in Chapter I, we are presently at the point where we are only beginning to identify in very general terms the types of forces which seem to play a role in accident causation. From the ideas and concepts put forth so far we can construct a generalized description of an accident model which might be applicable to the home accident problem. This was done in Chapter I. As indicated there, one side of the model consists of the human being, with an initial propensity to have accidents (or an average initial probability if viewed in the aggregate) which is shifted either upward or downward by the individual's interaction with an almost infinite variety of environmental forces (physical, socio-economic, and biological) of both long term and transient nature. On the other side of the model we have the idea that the environment can be more or less safe, strictly in terms of the actual opportunity to have accidents. Thus, leaving the relevant environmental forces strictly aside (or on the other side of the model), we might think in terms of a level of dangerous situations or hazards which may also vary in time and space. Although in a given situation the individual may be so "primed" for an accident that he is liable to have one in any physical setting, this is probably not the case in the majority of accidents. The individual may not need to have as high a propensity in order to actually have an accident if the physical setting is dangerous

enough. Thus the concept of a hazard may have meaning and we can think of the whole model in terms of an individual, acted on by environmental forces, and "in contact," so to speak, with a given level of environmental hazards.

Also very important is the idea of a distinction between the "mechanism" and the "agent" when speaking of environmental hazards. Currently available statistics deal in terms only of the agent, and these are complete enough so that we can construct tables of death rates by age group and agent (see Appendix Tables X through XII). Yet as useful as these might be in helping us pinpoint the specific age group-agent combinations in which the greatest numbers of deaths occur, we cannot mount preventive programs without knowledge of the types of mechanisms which are most important. This is where the data is lacking.

Turning to the analysis itself, perhaps the most important concept to be gained is the idea that the occurrence of individual home accidents, or home accident fatalities, within a given population tends to follow a bernoulli process, i.e., the individual faces a dichotomous situation with a certain probability p of an event occurring, and that the death rates in a group of counties with the same average probability p of a death occurring will tend to be distributed in binomial fashion.

Although we have very little information as to how most human and environmental variables affect this probability, one that we do have knowledge of is the effect of age. Thus another important facet of the analysis was the recognition of this and the use of a device called the "equivalent population" to take it into account. It was found, through the use of this device, that the age distribution of the population has substantial effects.

Finally, returning to the bernoulli process and the binomial distribution, it was recognized both intuitively and through examination of the standard deviation of the binomial distribution that the amount of information yielded by the data for an individual county varies with the population size of the county. It follows from this that more weight in the analysis should be given to the larger counties and thus the validity of a simple regression comes into question. No rigorous theoretical analysis was done here to attempt to determine what the weighting should be. Rather, it was assumed that weights should be distributed in proportion to population size.

Turning to the second theme, some evidence was found to indicate a relation between the percentage of dilapidated housing, as defined and enumerated by the U.S. Census, and the average home accident death rate, corrected for variations in the age distribution of county populations, for counties in the State of Kansas during the period from 1950 to 1957.

Several things should be kept in mind, however. The first of these is that the correlation obtained was not of a high level of statistical significance and it was obtained through the analysis of a limited set of data. Thus a great deal of importance should not be attached to it. Rather, it should be viewed as an indication that similar analyses of other sets of data might be worthwhile. We cannot begin to generalize until the analyses of several sets of data have repeatedly indicated that a relationship exists.

Further, this analysis says nothing about health in general as it relates to housing quality, nor does it say anything about home accident

injuries of all types. The analysis of data on all types of home accident injuries relative to housing quality would undoubtedly be more worthwhile than the analysis of fatality data only. Unfortunately, any such analysis will require a great deal of effort in data collection, since governmental agencies do not collect and publish such data. Although the U.S National Health Survey collects data on all types of injuries, it does not collect data on other variables which might be relevant to the causation of those injuries.

As was mentioned earlier, there are large gaps in our knowledge about most of the detailed relationships involved in the accident causation model. Much research remains to be done on all of its aspects. As a result of this fact it is impossible at present to attach a definite meaning to any correlation which may exist between home accident deaths and housing quality. Such a correlation might exist as a result of any one of or combination of the following factors:

- (a) Human factors or general environmental forces covariant with but unrelated to housing quality.
- (b) General environmental forces associated with housing quality.
- (c) Variation in the density of environmental hazards with housing quality.

Thus, as can readily be seen, even if a correlation can be consistently demonstrated with a variety of sets of data, all the questions will not be answered. These questions relate to the various elements of our accident model, i.e., the individual, the general environmental forces which may affect his propensity to have accidents, and the distribution of environmental hazards over the range of housing quality. The answers

will be found only if more research is done, using the generalized home accident model as a basis for the formulation of hypotheses.

APPENDIX

Table I*

Kansas Home Accident Deaths by County and Year, 1950-1957

County	1950	1951	1952	1953	1954	1955	1956	1957
Allen	6	8	--	2	2	5	2	7
Anderson	2	--	2	--	5	2	3	--
Atchinson	4	4	6	5	--	7	6	5
Barber	3	2	2	4	3	--	2	--
Barton	6	9	12	6	4	3	4	3
Bourbon	6	7	6	2	7	5	7	6
Brown	2	4	2	2	1	4	2	3
Butler	5	4	8	15	8	4	8	4
Chase	1	1	4	3	2	1	--	--
Chautauqua	1	1	1	1	6	1	4	2
Cherokee	10	4	6	2	7	7	9	3
Cheyenne	2	--	--	2	--	1	1	1
Clark	2	--	--	--	1	--	--	4
Clay	7	4	4	1	4	8	5	5
Cloud	4	6	1	4	5	1	7	4
Coffey	7	5	--	2	3	2	5	4
Comanche	1	2	4	--	2	1	1	1
Cowley	7	7	7	5	17	6	10	7
Crawford	13	19	10	6	29	13	14	7
Decatur	--	--	--	2	2	--	1	1
Dickinson	5	8	5	3	7	6	3	4
Doniphan	--	1	4	2	4	1	3	1
Douglas	10	5	7	5	5	8	4	7
Edwards	1	--	--	1	1	2	3	1
Elk	3	1	1	2	1	--	1	4
Ellis	--	5	2	5	--	4	1	6
Ellsworth	1	4	--	1	1	2	3	2
Finney	3	2	4	4	--	1	4	3
Ford	4	5	3	3	4	1	6	--
Franklin	4	4	9	5	4	2	3	8
Geary	6	5	4	7	3	8	6	7
Gove	1	--	1	1	--	--	1	1
Graham	1	--	--	2	--	1	1	1
Grant	1	--	1	--	1	--	--	1
Gray	1	1	2	--	1	--	3	4
Greeley	--	--	2	--	1	--	--	1
Greenwood	2	2	1	1	2	2	2	4
Hamilton	--	--	--	2	2	2	--	2
Harper	4	2	1	1	2	--	2	2
Harvey	4	4	2	1	2	3	1	5
Haskell	1	2	1	--	--	--	--	--
Hodgeman	1	--	--	--	--	1	--	1
Jackson	1	1	2	--	1	2	2	1
Jefferson	1	6	6	2	3	4	5	2

(Con't)

Table I (Con't)

County	1950	1951	1952	1953	1954	1955	1956	1957
Jewell	2	4	2	1	1	4	1	4
Johnson	9	17	7	8	11	9	13	8
Kearney	--	--	--	--	--	--	--	1
Kingman	2	1	3	2	1	--	1	2
Kiowa	--	1	--	--	--	--	1	--
Labette	8	10	12	5	12	10	7	5
Lane	--	--	--	--	--	1	--	--
Leavenworth	13	18	11	4	10	11	9	8
Lincoln	--	2	--	2	1	2	1	3
Linn	3	6	3	3	3	2	2	3
Logan	--	--	--	--	2	--	1	--
Lyon	6	3	9	7	10	2	9	3
Marion	3	5	2	2	6	5	1	5
Marshall	7	7	4	5	6	6	7	3
McPherson	5	1	4	5	5	1	3	4
Mead	1	--	1	2	--	1	--	1
Miami	8	5	5	2	3	2	4	4
Mitchell	1	3	1	5	3	2	1	1
Montgomery	16	17	10	10	18	9	11	11
Morris	4	3	2	1	4	--	1	1
Morton	1	1	--	3	1	--	--	--
Nemaha	1	4	3	1	3	2	4	1
Neosho	7	7	9	11	10	8	4	4
Ness	5	1	1	2	--	1	--	--
Norton	--	2	2	3	2	2	2	2
Osage	3	1	4	7	4	3	1	3
Osborne	5	2	--	2	5	4	1	1
Ottawa	2	--	2	2	2	4	1	--
Pawnee	4	4	1	6	3	2	1	2
Phillips	7	2	1	1	3	2	--	4
Pottawatomie	2	6	6	2	3	2	2	4
Pratt	4	--	6	3	--	5	1	2
Rawlins	3	1	1	1	--	1	1	1
Reno	16	8	11	6	13	10	9	8
Republic	4	4	2	1	5	2	2	1
Rice	2	2	7	1	3	4	2	3
Riley	8	10	7	1	3	6	6	4
Rooks	1	2	3	--	3	1	--	--
Rush	--	1	--	4	3	--	--	1
Russell	1	1	5	1	3	3	--	3
Saline	7	9	6	10	8	3	6	5
Scott	--	1	2	--	2	2	1	--
Sedgwick	46	50	44	29	44	37	47	42
Seward	4	2	3	4	1	2	4	1

(Con't)

Table I (Con't)

County	1950	1951	1952	1953	1954	1955	1956	1957
Shawnee	22	30	22	22	28	14	17	21
Sheridan	--	--	--	--	2	--	1	1
Sherman	--	--	--	--	--	2	--	2
Smith	3	6	2	1	1	6	1	4
Stafford	3	3	1	2	2	1	5	3
Stanton	--	--	1	--	--	--	--	--
Stevens	2	--	--	1	--	1	--	1
Sumner	8	11	5	5	7	5	3	2
Thomas	--	--	3	--	1	1	1	--
Trego	1	5	1	--	--	1	2	2
Wabaunsee	1	1	1	2	3	6	--	2
Wallace	--	--	1	--	--	1	--	--
Washington	6	4	2	2	7	1	2	1
Wichita	1	--	--	1	2	--	--	--
Wilson	--	1	6	1	1	1	5	2
Woodson	1	1	4	--	3	--	--	3
Wyandotte	50	39	35	29	45	40	21	32
Total	463	470	419	343	475	374	373	370

* Source: Kansas State Board of Health, Division of Vital Statistics, Kansas Accidental Death Report, biennial statistical report, covering the preceding two years (Kansas State Board of Health: 1952, 1954, 1956, 1958).

1952 edition (1950 & 1951) -- Table #39, pp. 90-92.
 1954 edition (1952 & 1953) -- Table #27, pp. 72-73.
 1956 edition (1954 & 1955) -- Table #28, pp. 92-93.
 1958 edition (1956 & 1957) -- Table #19, pp. 72-73.

Table II
Unadjusted Death Rates by Counties
(Home Accident Death Rates)

County	Population (10,000's)			Deaths- Eight Years	Deaths/ 10,000/ 8 years
	1950	1960	New		
Allen	1.82	1.64	1.75	32	18.33
Anderson	1.03	0.90	0.98	14	14.33
Atchison	2.15	2.09	2.13	37	17.41
Barber	0.85	0.87	0.86	15	17.45
Barton	2.99	3.24	3.09	47	15.21
Bourbon	1.92	1.61	1.79	46	25.66
Brown	1.47	1.32	1.41	20	14.20
Butler	3.10	3.84	3.40	56	16.49
Chase	0.48	0.39	0.45	12	26.86
Chautauqua	0.74	0.60	0.68	17	24.97
Cherokee	2.51	2.23	2.40	48	20.00
Cheyenne	0.57	0.47	0.53	7	13.25
Clark	0.39	0.34	0.37	7	18.79
Clay	1.17	1.07	1.13	38	33.66
Cloud	1.61	1.44	1.54	32	20.75
Coffey	1.04	0.84	0.96	28	29.15
Comanche	0.39	0.33	0.36	12	32.96
Cowley	3.69	3.79	3.73	66	17.70
Crawford	4.02	3.70	3.90	111	28.50
Decatur	0.62	0.58	0.60	6	9.97
Dickinson	2.12	2.16	2.13	41	19.21
Doniphan	1.05	0.96	1.01	16	15.80
Douglas	3.41	4.37	3.79	51	13.44
Edwards	0.59	0.51	0.56	9	16.05
Elk	0.67	0.50	0.60	13	21.57
Ellis	1.90	2.13	1.99	23	11.54
Ellsworth	0.85	0.77	0.82	14	17.12
Finney	1.51	1.61	1.55	21	13.56
Ford	1.97	2.09	2.02	26	12.89
Franklin	1.99	1.95	1.98	39	19.72
Geary	2.17	2.88	2.45	46	18.76
Gove	0.44	0.41	0.43	5	11.60
Graham	0.50	0.56	0.52	6	11.44
Grant	0.46	0.53	0.49	4	8.18
Gray	0.49	0.44	0.47	12	25.60
Greeley	0.20	0.21	0.20	4	19.60
Greenwood	1.36	1.13	1.26	16	12.65
Hamilton	0.37	0.31	0.35	8	23.02
Harper	1.03	0.95	1.00	14	14.04
Harvey	2.17	2.59	2.34	22	9.42
Haskell	0.26	0.30	0.28	4	14.49
Hodgeman	0.33	0.31	0.32	3	9.29

(Con't)

Table II (Con't)

County	Population (10,000's)			Deaths- Eight Years	Deaths/ 10,000/ 8 years
	1950	1960	New		
Jackson	1.11	1.03	1.08	10	9.27
Jefferson	1.11	1.13	1.12	29	26.01
Jewell	0.97	0.72	0.87	19	21.82
Johnson	6.28	14.38	9.52	82	8.61
Kearney	0.35	0.31	0.33	1	3.00
Kingman	1.03	1.00	1.02	12	11.79
Kiowa	0.47	0.46	0.47	2	4.26
Labotte	2.93	2.68	2.83	69	24.39
Lane	0.28	0.31	0.29	1	6.88
Leavenworth	4.24	4.85	4.48	84	18.74
Lincoln	0.66	0.56	0.62	11	17.72
Linn	1.01	0.83	0.93	25	26.76
Logan	0.42	0.40	0.41	3	7.25
Lyon	2.66	2.69	2.67	49	18.33
Marion	1.63	1.51	1.58	29	18.34
Marshall	1.79	1.56	1.70	45	26.48
McPherson	2.37	2.43	2.39	28	11.71
Meade	0.57	0.55	0.56	6	10.66
Miami	1.97	1.99	1.98	33	16.69
Mitchell	1.03	0.89	0.97	17	17.46
Montgomery	4.65	4.50	4.59	102	22.22
Morris	0.85	0.74	0.80	16	19.88
Morton	0.26	0.34	0.29	6	20.63
Nemaha	1.43	1.29	1.38	19	13.80
Neosho	2.03	1.95	2.00	60	30.01
Ness	0.63	0.55	0.60	10	16.72
Norton	0.88	0.80	0.85	15	17.65
Osage	1.28	1.29	1.28	26	20.25
Osborne	0.86	0.75	0.81	20	24.58
Ottawa	0.73	0.68	0.71	13	18.39
Pawnee	1.10	1.03	1.07	23	21.44
Phillips	0.93	0.87	0.90	20	22.11
Pottawatomie	1.23	1.20	1.22	27	22.15
Pratt	1.22	1.21	1.21	21	17.30
Rawlins	0.51	0.53	0.55	9	16.28
Reno	5.41	5.91	5.61	81	14.45
Republic	1.15	0.98	1.08	21	19.46
Rice	1.56	1.39	1.49	24	16.06
Riley	3.34	4.19	3.64	45	12.35
Rooks	0.90	0.97	0.93	10	10.73
Rush	0.72	0.62	0.68	9	13.25
Russell	1.34	1.13	1.26	17	13.51
Saline	3.34	5.47	4.19	54	12.88
Scott	0.49	0.52	0.50	8	15.86

(Con't)

Table II (Con't)

County	Population (10,000's)			Deaths- Eight Years	Deaths/ 10,000/ 8 years
	1950	1960	New		
Sedgwick	22.23	34.32	27.07	339	12.52
Seward	1.00	1.59	1.24	21	17.00
Shawnee	10.54	14.13	11.98	176	14.70
Sheridan	0.46	0.43	0.45	4	8.95
Sherman	0.74	0.67	0.71	4	5.64
Smith	0.88	0.78	0.84	24	28.52
Stafford	0.88	0.75	0.83	20	24.18
Stanton	0.23	0.21	0.22	1	4.54
Stevens	0.45	0.44	0.45	5	11.19
Sumner	2.36	2.53	2.43	46	18.92
Thomas	0.76	0.74	0.75	6	8.01
Trego	0.59	0.55	0.57	12	21.02
Wabaunsee	0.72	0.66	0.70	16	22.90
Wallace	0.25	0.21	0.23	2	8.58
Washington	1.30	1.07	1.21	25	20.69
Wichita	0.26	0.28	0.27	4	14.87
Wilson	1.48	1.31	1.41	17	12.04
Woodson	0.67	0.54	0.62	12	19.37
Wyandotte	16.53	18.55	17.34	291	16.78

Table III
Dilapidation Rate Calculations

County	1950		1960		Total		%
	O.U.	O.D.	O.U.	O.D.	O.U.	O.D.	
Allen	5989	433	5609	222	11598	655	5.65
Anderson	3309	198	2991	79	6300	277	4.40
Atchison	6311	793	6251	264	12562	1057	8.41
Barber	2715	137	2834	20	5549	157	2.83
Barton	8858	458	9846	219	18699	677	3.62
Bourbon	6416	674	5752	305	12168	979	8.05
Brown	4890	187	4549	213	9439	400	4.24
Butler	9787	771	11884	411	21671	1182	5.45
Chase	1568	108	1345	---	2913	108	3.71
Chautauqua	2456	290	2158	119	4614	409	8.86
Cherokee	8156	1152	7587	373	15743	1525	9.69
Cheyenne	1702	120	1505	20	3207	140	4.37
Clark	1290	111	1156	36	2386	147	6.16
Clay	3914	147	3708	115	7622	262	3.44
Cloud	5134	387	4736	321	9870	708	7.17
Coffey	3414	191	2928	50	6342	241	3.80
Comanche	1232	74	1130	17	2362	91	3.85
Cowley	11507	987	12207	210	23264	1197	5.15
Crawford	13587	1100	13085	560	26672	1660	6.22
Decatur	2017	146	1904	25	3921	171	4.36
Dickinson	6842	331	7110	263	13952	594	4.26
Doniphan	3290	302	3002	167	6292	469	7.45
Douglas	9358	465	12392	514	21750	979	4.50
Edwards	2054	51	1674	88	3728	139	3.73
Elk	2229	400	1846	97	4075	497	12.20
Ellis	4838	168	5692	55	10531	223	2.12
Ellsworth	2653	338	2545	224	5198	562	10.81
Finney	4350	470	4650	208	9020	678	7.53
Ford	6027	386	6386	192	12413	578	4.66
Franklin	6380	324	6546	210	12926	534	4.13
Geary	6071	268	7758	147	13829	415	3.00
Gove	1176	64	1167	37	2343	101	4.31
Graham	1470	142	1664	37	3134	179	5.71
Grant	1274	89	1447	31	2721	120	4.41
Gray	1397	71	1337	21	2734	92	3.37
Greeley	567	39	619	58	1186	97	8.18
Greenwood	4382	161	3927	217	8309	378	4.55
Hamilton	1080	87	963	49	2043	136	6.66
Harper	3335	188	3272	36	6607	224	3.39
Harvey	6605	379	7986	401	14591	780	5.35
Haskell	745	52	848	15	1593	67	4.21
Hodgeman	918	29	930	26	1848	55	2.98

(Con't)

Table III (Con't)

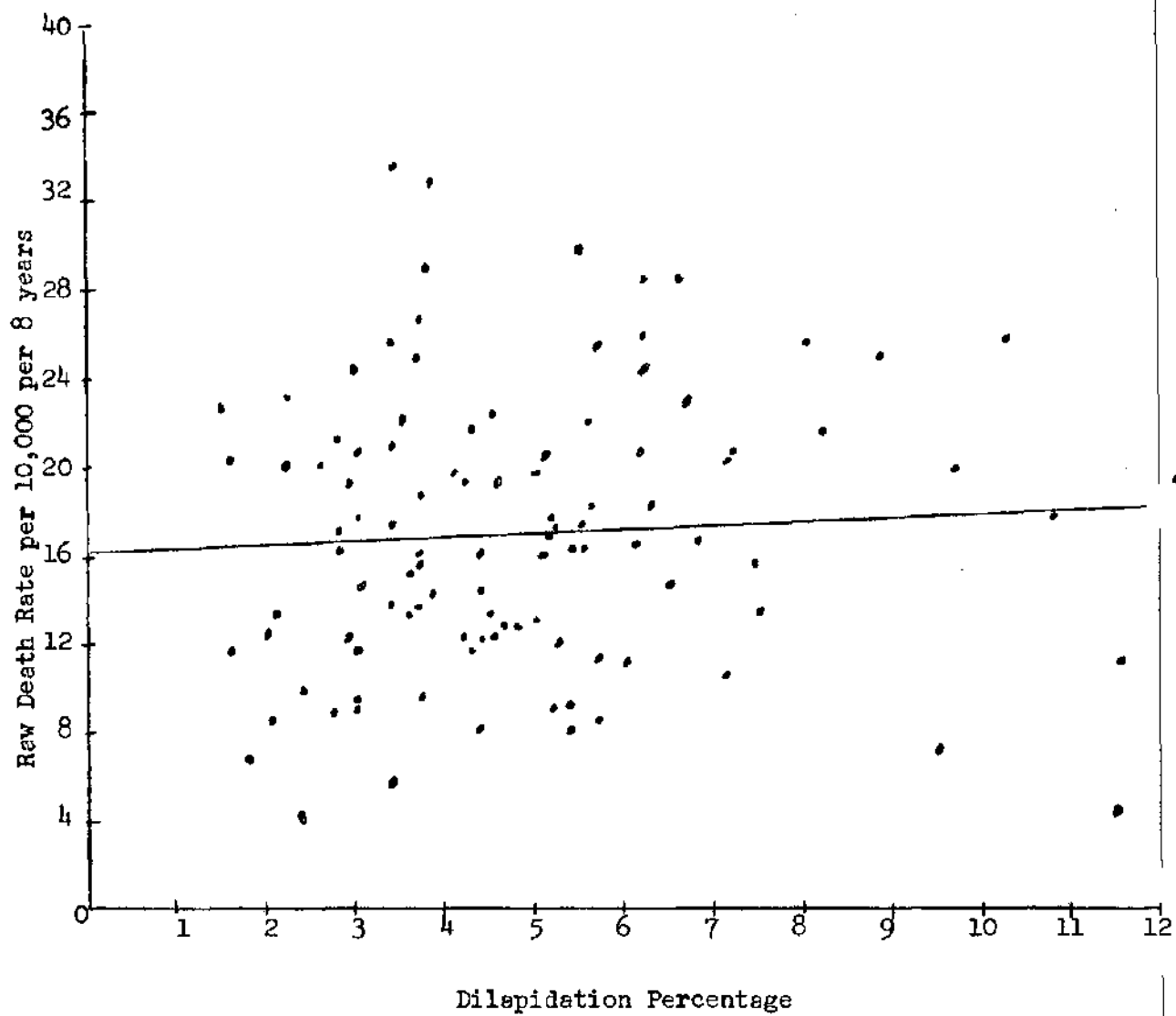
County	1950		1960		Total		%
	O.U.	O.D.	O.U.	O.D.	O.U.	O.D.	
Jackson	3608	176	3426	188	7034	364	5.17
Jefferson	3530	278	3473	157	7003	435	6.21
Jewell	3158	182	2532	64	5690	246	4.32
Johnson	18751	640	40270	578	59021	1218	2.06
Kearney	964	137	895	40	1859	177	9.52
Kingman	3156	51	----	---	3156	51	1.62
Kiowa	1487	15	1469	57	2956	72	2.44
Labette	9348	705	8971	420	18319	1125	6.14
Lane	806	9	930	22	1736	31	1.79
Leavenworth	10991	899	12394	780	23385	1679	7.18
Lincoln	2143	94	1906	28	4049	122	3.01
Linn	3355	500	2944	149	6299	649	10.30
Logan	1208	98	1206	131	2414	229	9.49
Lyon	8391	509	8492	203	16883	712	4.22
Marion	5020	129	4818	85	9838	214	2.18
Marshall	5807	422	5263	206	11070	628	5.67
McPherson	7270	272	7551	170	14821	440	2.98
Meade	1673	107	1726	20	3399	127	3.74
Miami	5732	405	5996	314	11728	719	6.13
Mitchell	3188	186	2930	135	6118	321	5.25
Montgomery	15430	1026	15342	706	30754	1732	5.63
Morris	2761	150	2468	110	5229	260	4.97
Morton	799	10	987	18	1786	28	1.57
Nemaha	4246	224	3891	75	8137	299	3.67
Neosho	6635	441	6437	278	13072	719	5.50
Ness	1915	118	1751	82	3666	200	5.46
Norton	2721	116	2645	66	5366	182	3.39
Osage	4006	110	4227	101	8233	211	2.56
Osborne	2786	47	2539	111	5325	158	2.97
Ottawa	2429	232	2347	69	4776	301	6.30
Pawnee	2753	41	3033	119	5786	160	2.77
Phillips	3018	233	2909	36	5927	269	4.54
Pottawatomie	3796	204	3730	58	7526	262	3.48
Prairie	3904	264	4050	174	7954	438	5.51
Rawlins	1631	56	1554	34	3185	90	2.83
Reno	17061	863	18916	524	35997	1387	3.86
Republic	3840	150	3420	63	7260	213	2.93
Rice	4964	326	4517	156	9481	482	5.08
Riley	8104	323	10714	225	18818	548	2.91
Rooks	2746	380	2958	27	5704	407	7.14
Rush	2180	145	2032	65	4212	210	4.99
Russell	4140	121	3678	158	7818	279	3.57
	10522	527	16201	757	26723	1284	4.80
	1416	57	1554	52	2970	109	3.67

(Con't)

Table III (Con't)

County	1950		1960		Total		%
	O.U.	O.D.	O.U.	O.D.	O.U.	O.D.	
Sedgwick	69426	3259	103422	235	172848	3494	2.02
Seward	3022	233	4700	164	7722	397	5.14
Shawnee	32774	1854	43625	480	76399	2334	3.06
Sheridan	1263	54	1211	14	2474	68	2.75
Sherman	2215	132	2154	17	4369	149	3.41
Smith	2949	309	2698	63	5647	372	6.59
Stafford	2838	140	2576	61	5414	201	3.71
Stanton	506	80	599	47	1105	127	11.49
Stevens	1298	63	1324	95	2622	158	6.03
Sumner	7701	425	8222	171	15923	596	3.14
Thomas	2258	86	2267	160	4525	246	5.44
Trego	1687	29	1630	84	3317	113	3.41
Wabaunsee	2310	40	2182	27	4492	67	1.49
Wallace	720	53	603	22	1323	75	5.67
Washington	4106	309	3552	81	7658	390	5.09
Wichita	737	28	788	71	1525	99	6.49
Wilson	4910	278	4490	216	9400	494	5.26
Woodson	248	111	1912	78	4130	189	4.58
Wyandotte	49245	4522	55872	2629	105117	7151	6.80

Figure I
Graph of Raw Data



Regression equation:

$$y = 16.12 + 0.174x$$

Table IV
 Theoretical Binomial Distribution
 $n = 3000, p = .0016$

u	Rate/10000	Pr
0	0	.008
1	3.33	.039
2	6.67	.094
3	10.00	.151
4	13.33	.183
5	16.67	.174
6	20.00	.139
7	23.33	.095
8	26.67	.057
9	30.00	.035
10	33.33	.016
11	36.67	.007

Using Pr. = $\frac{n!}{u!(n-u)!} p^u q^{(n-u)}$

Table v
 Theoretical Binomial Distribution
 n = 10000, p = .0016

u	Rate/10000	Pr
0	0	--
1	1	--
2	2	--
3	3	--
4	4	--
5	5	.0009
6	6	.0026
7	7	.0059
8	8	.0116
9	9	.0210
10	10	.0338
11	11	.0490
12	12	.0654
13	13	.0808
14	14	.0920
15	15	.0985
16	16	.0985
17	17	.0925
18	18	.0822
19	19	.0694
20	20	.0555
21	21	.0423
22	22	.0307
23	23	.0214
24	24	.0142
25	25	.0092
26	26	.0056
27	27	.0033
28	28	.0019
29	29	.0011
30	30	.0006

Using $Pr = \frac{n!}{u!(n-u)!} p^u q^{(n-u)}$

Table VI

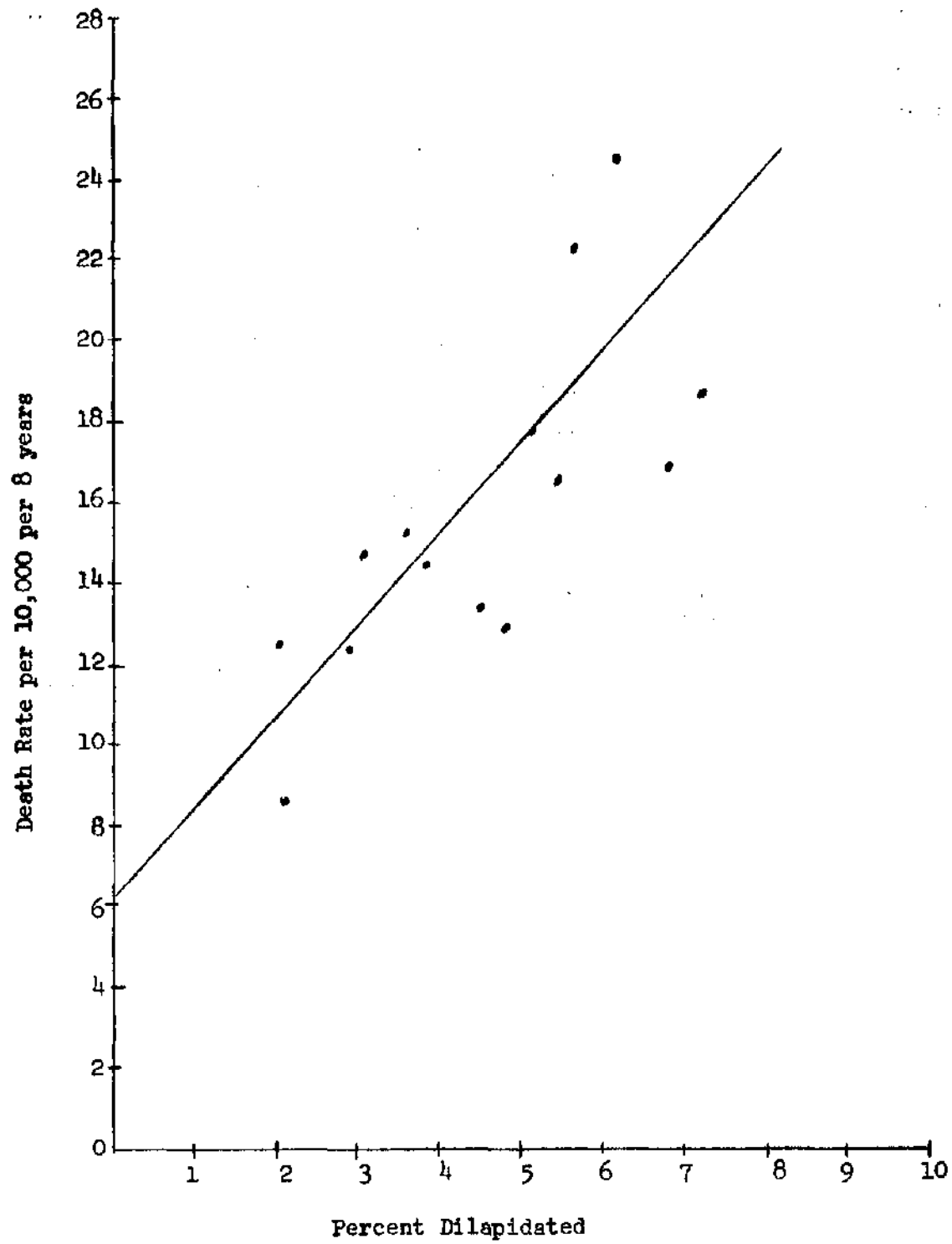
Actual Distribution:
 Set of 41 Counties with $p = .00163$
 Average % Dilapidation 3%-5%

Rate/10000	Mean Rate	#	%/100	Adjusted (Arbitrary)
0-1.67	0.83	0	.00	.00
1.68-5.00	3.33	0	.00	.00
5.01-8.33	6.67	2	.049	.049
8.34-11.67	10.00	3	.073	.120
11.68-15.00	13.33	14	.342	.22
15.01-18.33	16.67	6	.146	.22
18.34-21.67	20.00	7	.170	.17
21.68-25.00	23.33	4	.098	.098
25.01-28.33	26.67	2	.049	.049
28.34-31.67	30.00	1	.024	.024
31.68-35.00	33.33	2	.049	.049
35.01-38.33	36.67	0	.00	.00

Table VII
Data on Sixteen Largest Counties

County	% Dil.	Pop.	Death Rate	Eq. Pop.	Death Rate
Sedgwick	2.02	27.06	12.5	21.70	15.4
Johnson	2.06	9.53	8.6	7.12	11.6
Riley	2.91	3.69	12.4	2.85	15.8
Shawnee	3.06	11.98	14.7	11.68	15.0
Barton	3.62	3.08	15.2	2.51	18.8
Reno	3.86	5.61	14.5	5.47	14.8
Lyon	4.22	2.67	18.3	3.03	16.1
Douglas	4.50	3.79	13.4	3.46	12.7
Saline	4.80	4.19	12.9	3.76	14.4
Cowley	5.15	3.73	17.7	4.22	15.8
Butler	5.45	3.40	16.5	3.28	17.0
Montgomery	5.63	4.59	22.2	5.19	19.6
Labette	6.14	2.83	24.4	3.42	20.2
Crawford	6.22	3.90	28.5	4.78	23.2
Wyandotte	6.80	17.34	16.8	15.23	19.1
Leavenworth	7.18	4.48	18.7	4.25	19.8

Figure II
16 Counties - Unadjusted



Regression equation: $y = 6.15 + 2.29x$

Figure III
16 Counties - Adjusted

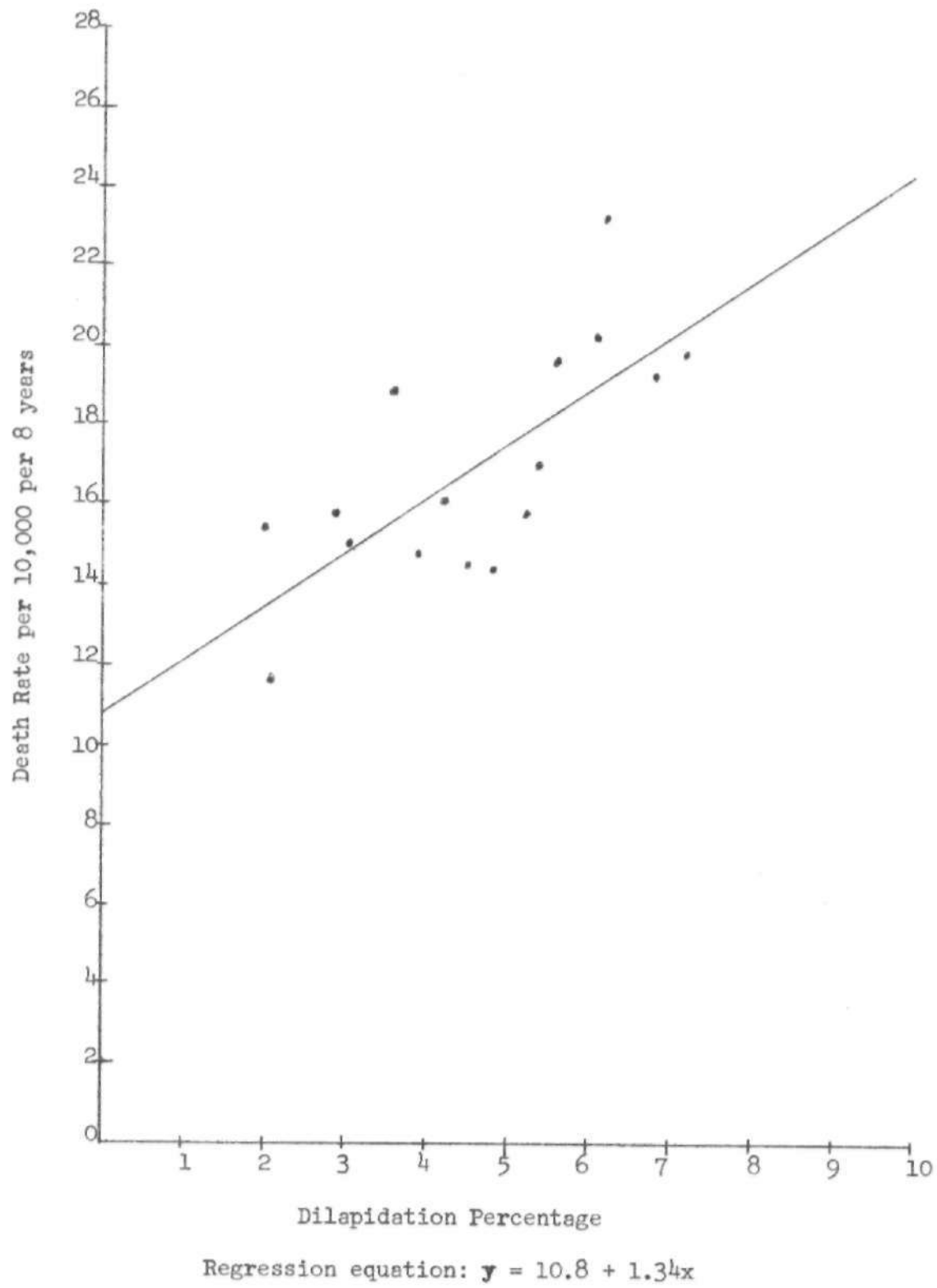


Table VIII-A
Counties with % Dilapidation 0-2.99

County	OU	OD	Pop.	Deaths
Barber	5549	157	8598	15
Ellis	10531	223	19934	23
Hodgeman	1848	55	3232	3
Johnson	59021	1218	95187	82
Kingman	3156	51	10178	12
Kiowa	2956	72	4695	2
Lane	1736	31	2909	1
Marion	9838	214	15841	29
McPherson	14821	440	23916	28
Morton	1786	228	2908	6
Osage	8233	211	12841	26
Osborne	5325	158	8137	20
Pawnee	5786	160	10726	23
Rawlins	3185	90	5528	9
Republic	7260	213	10794	21
Riley	18818	548	36449	45
Sedgwick	172848	3494	270666	339
Sheridan	2474	68	4471	4
Wabaunsee	4492	67	6986	16
Total	339663	7498	553996	704

$$\% \text{ Dilapidation} = \frac{7498}{339663} = 2.21\%$$

$$\text{Deaths per 10,000 per 8 years} = \frac{704}{55.4} = 12.71$$

Table VIII-B
Counties with % Dilapidation 3.00-4.99

County	OU	OD	Pop.	Deaths
Anderson	6300	277	9772	14
Barton	18699	677	30893	47
Brown	9439	400	14082	20
Chase	2913	108	4468	12
Cheyenne	3207	140	5284	7
Clay	7622	262	11228	38
Coffey	6342	241	9606	28
Comanche	2362	91	3641	12
Decatur	3921	171	6021	6
Dickinson	13952	594	21343	41
Douglas	21750	979	37940	51
Edwards	3728	139	5609	9
Ford	12413	578	20177	26
Franklin	12926	534	19776	39
Geary	13829	415	24514	46
Gore	2343	101	4311	5
Grant	2721	120	4890	4
Gray	2734	92	4688	12
Greenwood	8309	378	12646	16
Harper	6607	224	9974	14
Haskell	1593	67	2760	4
Jewell	5690	246	8709	19
Lincoln	4049	122	6208	11
Lyon	16883	712	26738	49
Meade	3399	127	5628	6
Morris	5229	260	8048	16
Nemaha	8137	299	13764	19
Norton	5366	182	8499	15
Phillips	5927	269	9047	20
Pottawatomie	7526	262	12189	27
Reno	35997	1387	56057	81
Rush	4212	210	6795	9
Russell	7818	279	12583	17
Saline	26723	1284	41931	54
Scott	2970	109	5044	8
Shawnee	76399	2334	119765	176
Sherman	4369	149	7097	4
Stafford	5414	201	8270	20
Sumner	15923	596	24314	46
Trego	3317	113	5710	12
Woodson	4130	189	6196	12
Totals	413078	15918	656215	1072

% Dilapidation = 3.85%

Death Rate = 16.34 per 10,000 population per 8 years.

Table VIII-C
 Counties with % Dilapidation 5.00-6.99

County	CU	OD	Pop.	Deaths
Allen	11598	655	17460	32
Butler	21671	1182	33958	56
Clark	2386	147	3726	7
Cowley	23264	1197	37287	66
Crawford	26672	1660	38951	111
Graham	3134	179	5246	6
Hamilton	2043	136	3475	8
Harvey	14591	780	23365	22
Jackson	7034	364	10782	10
Jefferson	7003	435	11151	29
Labette	18319	1125	28293	69
Marshall	11070	628	16995	45
Miami	11728	719	19772	33
Mitchell	6118	321	9738	17
Montgomery	30754	1732	45895	102
Neosho	13072	719	19992	60
Ness	3666	200	5981	10
Ottawa	4776	301	7069	13
Pratt	7954	438	12142	21
Rice	9481	482	14945	24
Seward	7722	397	12355	21
Smith	5647	372	8418	24
Stevens	2622	158	4470	5
Thomas	4525	246	7486	6
Wallace	1323	75	2332	2
Washington	7658	390	12082	25
Wichita	1525	999	2690	4
Wilson	9400	494	14120	17
Wyandotte	105117	7151	173389	291
Total	381819	22782	603565	1136

% Dilapidation = 5.97%

Death Rate = 18.82 per 10000 population per 8 years.

Table VIII-D
Counties with % Dilapidation 7.00 and Over

County	OU	OD	Pop.	Deaths
Atchison	12562	1057	21257	37
Bourbon	12168	979	17928	46
Chautauqua	4614	409	6808	17
Cherokee	15743	1525	23998	48
Cloud	9870	708	15425	32
Doniphan	6292	469	10129	16
Elk	4075	497	6027	13
Ellsworth	5198	562	8150	14
Finney	9000	678	15492	21
Greeley	1186	97	2041	4
Kearney	1859	177	3338	1
Linn	6299	649	9341	25
Logan	2414	229	4138	3
Rooks	5704	407	9319	10
Stanton	1105	127	2201	1
Leavenworth	23385	1679	44826	84
Total	121474	10249	200418	372

% Dilapidation = 8.44

Death Rate = 18.56 per 10,000 population per 8 years.

Table IX
Equivalent Population Calculations

Death Rate Ratios:			
Age Group	Rate	Avg. Rate	Ratio
0-4	20.5	16.34	1.255
5-44	3.68	16.34	0.225
45-64	8.35	16.34	0.51
65+	97.1	16.34	5.93

County Groups:			
<u>Group I</u>			
Age Group	Pop.	Ratio	Eq. Pop.
0-4	6.64	1.255	8.3
5-44	34.14	0.225	7.7
45-64	10.17	0.51	5.2
65 and over	4.48	5.93	26.7
Total	55.4	----	47.9

<u>Group II</u>			
Age Group	Pop.	Ratio	Eq. Pop.
0-4	6.93	1.255	8.8
5-44	37.69	0.225	8.5
45-64	13.56	0.51	6.9
65 and over	7.32	5.93	43.5
Total	65.6	----	67.7

(Con't)

Table IX. (Con't)

Group III

Age Group	Pop.	Ratio	Eq. Pop.
0-4	6.20	1.255	7.8
5-44	33.98	.225	7.7
45-64	13.13	.51	6.7
65 and over	7.04	5.93	41.7
Total	60.4	--	63.9

Group IV

Age Group	Pop.	Ratio	Eq. Pop.
0-4	2.01	1.255	2.5
5-44	11.28	.225	2.5
45-64	4.33	.51	2.2
65 and over	2.43	5.93	14.4
Total	20.0	--	21.6

Table X*
 Distribution of Home Accident Deaths by Type and Age Group,
 Kansas, 1950-1957

Accident Type	Age Group						Total
	0-4	5-14	15-24	25-44	45-64	65+	
Falls	20	9	3	23	85	1585	1725
Burns, Explosions, Conflagrations	139	43	26	77	111	268	666
Poisonous Gas	5	6	23	31	44	44	153
Firearms	11	35	23	32	30	6	137
Poisoning--Liquid or Solid	37	2	4	15	24	14	96
Mech. Suffocation	102	9	0	1	0	2	114
Suffocation, Foreign Object	73	2	0	0	3	4	82
Drownings	16	3	3	3	6	8	39
All Others	45	19	3	34	41	133	275
	448	128	85	216	344	2064	3287

* Source: Kansas State Board of Health, Division of Vital Statistics,
 Kansas Accidental Death Report.

1952 edition -- Table #33, p. 78, and Table #34, p. 81.

1954 edition -- Table #22, p. 62, and Table #23, p. 65.

1956 edition -- Table #24, p. 82, and Table #25, p. 86.

1958 edition -- Table #15, p. 62, and Table #16, p. 66.

Table XI
 Percentage Distribution of Home Accident Deaths
 by Type and Age Group, Kansas, 1950-1957

Accident Type	Age Group						Total
	0-4	5-14	15-24	25-44	45-64	65+	
Falls	0.6	0.3	0.1	0.7	2.6	48.2	52.5
Burns, Explosions, Conflagrations	4.2	1.3	0.8	2.3	3.4	8.2	20.2
Poisonous Gas	0.2	0.2	0.7	1.0	1.3	1.3	4.7
Firearms	0.3	1.1	0.7	1.0	0.9	0.2	4.2
Poisoning--Liquid or Solid	1.1	0.1	0.1	0.5	0.7	0.4	2.9
Mech. Suffocation	3.1	0.3	0.0	0.0	0.0	0.1	3.5
Suffocation, Foreign Object	2.2	0.1	0.0	0.0	0.1	0.1	2.5
Drownings	0.5	0.1	0.1	0.1	0.2	0.2	1.2
All Others	1.4	0.6	0.1	1.0	1.2	4.0	8.3
Total	13.6	3.9	2.6	6.6	10.4	62.7	100.0

Table XII
 Home Accident Death Rates Per 100,000 Population Per Year
 by Type and Age Group, Kansas, 1950-1957

Accident Type	Age Group						Total
	0-4	5-14	15-24	25-44	45-64	65+	
Falls	1.15	0.33	0.13	0.54	2.58	93.2	10.70
Burns, Explosions, Conflagrations	7.05	1.58	1.16	1.80	3.37	15.8	4.12
Poisonous Gas	0.29	0.22	1.02	0.72	1.33	2.58	0.95
Firearms	0.63	1.28	1.02	0.75	0.91	0.35	0.85
Poisoning--Liquid or Solid	2.13	0.07	0.18	0.35	0.73	0.83	0.60
Mech. Suffocation	5.87	0.33	0.00	0.02	0.00	0.12	0.71
Suffocation, Foreign Object	4.20	0.07	0.00	0.00	0.09	0.24	0.51
Drownings	0.92	0.11	0.13	0.07	0.18	0.47	0.24
All Others	2.58	0.70	0.13	0.80	1.24	7.83	1.70
Total	25.8	4.72	3.78	5.05	10.45	121.5	20.4

BIBLIOGRAPHY

- American Medical Association, Impact of Environment on Accidental Injuries and Fatalities, Proceedings, Third AMA Congress on Environmental Health Problems, April 4-5, 1966 (Chicago: 1966).
- Azar, Gordon J., and Alfred H. Lawton, "Sensory and Perceptual Changes that May Influence Housing Needs of the Aging," Patterns of Living and Housing of Middle-Aged and Older People (U.S. Department of Health, Education, and Welfare, Public Health Service: 1965), pp. 11-15.
- Azar, Gordon J., and Alfred H. Lawton, "Consequences of Physical and Physiological Change with Age in the Patterns of Living and Housing for the Middle-Aged and Aged," Patterns of Living and Housing of Middle-Aged and Older People (U.S. Department of Health, Education, and Welfare, Public Health Service, 1965), pp. 19-26.
- Croxton, Frederick E., Dudley, J. Cowden and Sidney Klein, Applied General Statistics (Englewood Cliffs, N.J.: Prentice-Hall, 1967).
- Gordon, John E., "The Epidemiology of Accidents," American Journal of Public Health, Volume 39 (April 1949), pp. 505-515.
- Gray, Thomas H., and George V. Truss, "Growth of an Accident Prevention Program," Public Health Reports, Volume 73 (June 1958), pp. 493-498.
- Hall, Edward T., The Hidden Dimension (Garden City, New York: Doubleday, 1966).
- Kansas State Board of Health, Division of Vital Statistics, Kansas Accidental Death Report (Kansas State Board of Health, biennial: 1952, 1954, 1956, 1958).