# LOG-LINEAR MODELS FOR EVALUATING HUNTING DEMAND.

# LOG-LINEAR MODELS FOR EVALUATING HUNTING DEMAND

by

Patricia Marie O'Neil

---

A Thesis Submitted to the Faculty of the

DEPARTMENT OF AGRICULTURAL ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1983

STATEMENT BY AUTHOR

This thesis has been submitted in partial fullfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

SIGNED: Patricia Marie O'Neil

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

William E. Martin
WILLIAM E. MARTIN
Professor of Agricultural Economics

July 27, 83
Date

ACKNOWLEDGEMENTS

I offer my sincere gratitude to the following people for their contributions in completion of this thesis: Dr. William E. Martin for instruction and support through my studies, guidance in the research and editing the thesis; Mr. David Smith, Mr. Gary Edwards, and Ms. Shirley Porterfield for their help in data collection and assimillation; Ms. Susan Ciolek-Torrello for her help in the computer work necessary for the research; Dr. Edwin Carpenter and Dr. David Barkley for their economic consultation; Ms. Sheila O'Connell and Ms. Ruby Hunnicutt for repeatedly typing preliminary drafts; and Ms. Peaches LaRue for her love and unselfish support.

TABLE OF CONTENTS

TABLE OF CONTENTS—*Continued*

LIST OF ILLUSTRATIONS

LIST OF TABLES

## ABSTRACT

Valuation of recreation resources is essential to formulation of policy decisions which reflect an efficient allocation of resources among alternative uses. For resource values to be ascertained, demand for the resources must be estimated. In this study, the log-linear model is evaluated as a method for determining hunting demand for deer and javelina at hunting sites in Arizona.

In the hypothesized model, distance is taken as a surrogate for cost. The objective is to test for site quality differences at given travel distances. Individudal observations of hunt applicants provide the data for the three decisions variables tested in the analysis the number of miles traveled to the hunting site, the probability of an applicant drawing a hunting permit to the site, and the probability of killing an animal at the site.

CHAPTER 1

# INTRODUCTION

## Problem Setting

The value of outdoor recreation concerns governments because one of their functions is the allocation of funds and resources among competitive uses. For many recreation resources it is infeasible to collect a charge. Assessments of value occurs through observation of public action rather than through private bids. In Arizona, outdoor recreation opportunities generally occur on public lands. Federal and state agencies have the management responsibilities. In order to allocate the resources expended on outdoor recreation in an economically rational manner, these agencies need valuation methods that quantify and compare the benefits associated with proposed projects.

Market prices play a dominant role in the allocation of most resources. However, the market price may not represent the total value that a good provides for an individual or for a society (Sublette and

> Market demand values tend to be good indicators of a resource allocation where the great majority of the benefit received is individual in character and is fully recognized by that consumer with respect to both immediate and future effects . . . . Great art, education, hospitals, roads and many other so-called 'community services' fall more or less into the same category as outdoor recreation, wherein, many social values to the individual are involved, often having long term implications for the future of society.

Martin, 1975). Beazley feels that because of this fact the market demand value for outdoor recreation may be an inappropriate measure of value in resource allocation (Beazley, 1961, p. 647).

Planners and administrators, until recently, have been unwilling to incorporate monetary values for recreation resources into the management decisions that effect resource use. Perhaps they felt that the intangible benefits Beazley speaks of were incapable of being measured. In essence, recreation resources were considered priceless. But the issue relating to recreation resources is not one of absence of value but absence of value measurement and comparison (Sublette and Martin, 1975). One purpose of this paper is to contribute to improvements in the methods of nonmarket valuation.

In July of 1978, President Carter published the directive "Improvement in the Planning and Evaluation of Federal Water Resource Programs and Projects." The President noted that deficiencies existed in methods of estimating benefits and costs of environmental and recreational resources. Direction was given to the Water Resource Council to develop a planning manual and define principles and standards to be used in evaluating federal projects.

The Water Resource Council first published its procedures in May of 1979. In a December 14, 1979 revision of "Principles and Standards" in the Federal Register (WRC, 1979) rules pertaining to "Recreation Valuation Methods and Procedures" were moved to the appendix and considered not binding. WRC viewed recreation valuation methods and procedures as still requiring refinement before they could be considered to hold rule status (WRC, 1979; p. 72895).

The need to address significant conceptual methodological and data collection problems associated with recreational demand has been recognized by the formation of a cooperative regional project (#W-133) entitled "Outdoor Recreation and the Public Interest: Evaluation of Benefits and Costs in Federal and State Resource Planning." There is no concerted regional effort elsewhere to test, apply, and evaluate recreation valuation methods and procedures. The project is designed to study the development of western resources, and to contribute conceptual and empirical knowledge to the federal evaluation procedures on a regional basis. This thesis is a contribution to Regional Research Project W-133. Its purpose is to extend the development of recreation demand methods to include quality comparisons among hunting sites. Because much of the land in Arizona is federal or state owned and operated, the refinement of public evaluation procedures are of particular importance to the state's citizens. The current Reagan administration has decreased the importance of specific recreation evaluation procedures. However, researchers continue their search for improvements, not only for knowledge alone, but in the hope that recreation valuation may be useful to governments at some future date.

## Problem Definition and Tactics

Most recent work in recreational demand has examined the recreation experience by one of several methods discussed in the next chapter. What these methods have in common is that they estimate the demand for activities that are assumed to be of homogeneous quality. Demand may be estimated for an entire region containing many recreation sites or for a particular site, but the quality of the site or sites has not been included as a variable in the demand equation. In Arizona, where diverse climate and terrain distinguish hunting sites, the assumption of homogeneity may not be relevant. This study seeks to incorporate decision variables indicative of site quality into the demand analysis and draw conclusions about the relationship between site quality and quantity demanded.

The general objective of this paper is to test the applicability of using a log-linear analysis of probabilistic data to model hunters' decisions to demand alternative quality sites. The specific objectives of this paper and an outline of the procedures used to accomplish these objectives follow.

Objectives

1. Review and compare valuation techniques.
2. Determine the significant decision variables pertinent to hunting demand for deer and javelina in the hunting season 1980-1981.
3. Determine the observed psuedo-demand function for each quality of hunting site.
4. Determine the fitted psuedo-demand function for each quality of hunting site.
5. Compare differences in site demand due to differences in the quality of the experience.

Procedure for Objective (1)

The easiest way to compare the value of goods and services is with market prices. When market prices are absent, alternative valuation methods are necessary. In order to assess and compare resource values the demand for the resource must first be estimated. Two methods of nonmarket good demand estimation, the travel cost method and the willingness to pay approach are reviewed. From demand estimates, resource values can be ascertained and the welfare positions of consumers examined. Consumers' surplus and the nondiscriminating monopolist approaches to resource valuation are reviewed. Finally, the log-linear model is reviewed in general as a statistical model and then specifically in terms of how it can be used to model hunting demand and valuation.

Following the empirical analysis, the log-linear method of estimating psuedo-demand functions is compared with the more traditional techniques of demand estimation for applicability of use in recreation valuation.

Procedure for Objective (2)

In order to determine the significant decision variables in hunting demand it is necessary to make hypotheses about hunters' behavior. First, since hunting permits are limited, the hunter must consider the odds that he will receive a permit for the given hunt for which he has applied. The odds of receiving a permit for a particular hunt are determined by the availability of permits and the level of demand for that hunt. The hunter could utilize information published by the Arizona Game and Fish Department (Arizona Game Survey and Harvest Data Summary) for the previous season's probabilities in determining his odds. It is hypothesized that the first decision variable in determining site demand is the "Probability of Draw." This is the probability of a hunter being drawn and permitted for his first choice of hunt.

Secondly, it is hypothesized that a hunter will consider the odds of successfully killing an animal on a particular hunt. These odds are determined by the hunters proficiency, the availability of animals at the site and the number of hunters at the site. The hunter could base hsi decision on published probabilities calculated from the previous season's hunts (Arizona Game Survey and Harvest Data Summary). The second decision variable, the "Probability of Success" of kill serves as a proxy commodity or benefit from the recreation experience from which the hunter receives utility.

Lastly, the hunter must consider the cost of his excursion. The location of the hunting site relative to the hunter's residence or the "miles traveled" is used as a proxy cost decision variable. It is expected that the hunter's decision to apply for a particular hunt and thus demand a particular site is based on the conclusions he draws about the relative importance of the site location, the probability of draw, and the probability of success.

Given these hypotheses, individual observations of the three statistics are obtained from hunters' permit applications to the Arizona Game and Fish Department for the season 1980-1981. These data are arranged in a three-dimensional cross-categorical continency table by the log-linear program. Each axis of the table corresponds to a decision variable. The axes are subdivided into reasonable intervals or categories over which the observed variable value may fall. For example, the probability of draw ($P_d$) may be categorized as low, medium, or high. Observed values of $P_d$ = 0.05, 0.1, 0.2 may fall in the low category. An individual hunter's observation is identified by a set of characteristics of the three variables. This observation is recorded in a corresponding cell of the three-way contingency table. Once all hunters of a particular animal have been categorized, the log-linear analysis determines which decision variables or interactions among the variables are significant in determining demand. The method of testing the hypotheses and levels of significance is discussed in the next section.

Procedure for Objective (3)

Once the contingency table has been established, the observed psuedo-demand functions can simply be read from the cells of the table. Each cell of the table represents some combination of probability of draw ($P_d$) and probability of success ($P_s$) over several intervals of miles traveled (M). A specific combination of $P_s$ and $P_d$ defines one or several hunting sites of a given quality. The spectrum of observations in a cell then represents the number of applicants demanding to travel various distances to a certain quality site. Each cell in the table represents a discontinuous step function which is the observed psuedo-demand function for a certain quality site. Where traditional demand functions relate quantity to price, the psuedo-demand function relates quantity of applicants, as determined by the odds of receiving benefit ($P_d$ and $P_s$), to the cost incurred (M) in receiving benefit.

Procedure for Objective (4)

The log-linear analysis tests all possible combinations of the decision variables for significance. These combinations or models range from the most simple single variable model to the most complex three way interaction. By the examination of chi square and maximum liklihood statistics associated with each model, the most appropriate model can be chosen. The log-linear program will fit the chosen model to the data and provide a contingency table of expected cell frequencies. The observations in each cell of the table represent the fitted psuedo-demand function for sites of equal quality.

The fitted function relates the overall relationship between the number of applicants for various quality sites to the miles traveled to those sites. This proxy quantity to cost relationship then may be used to infer the marginal utility of hunting at a particular quality of site gained from traveling an additional number of miles.

Procedure for Objective (5)

A comparison of quality between sites is obtained graphically as the vertical distance between fitted psuedo-demand functions. Quality is denoted by the number of additional votes made by applicants toward one site as opposed to another, where both sites lie equal distances from the voters. A difference in the quantity of applicants at two sites of equal mileages infers utility gained from differences in the quality variables $P_s$ and $P_d$.

Policy decisions for game management generally effect the probabilities of draw and success either by altering animal populations or by adjusting the availability of hunting permits. Inferences about how hunters value these quality variables can directly aid in resource planning and policy decisions.

# CHAPTER 2

# ECONOMIC DEMAND

## Conventional Demand Curves

Economic values are measured by what people are willing to give up in order to enjoy possession of a good or service. This explanation of value is conceptually the same for an outdoor recreation experience as it is for any other good or service, except that most outdoor recreation goods lack a formal market-determined price. Consumers must receive satisfaction or utility equal to the maximum price that they are willing to incur for a good. To act otherwise would mean to incur expense irrationally. With a market-price commodity, the amount of the product that a person purchases is regulated by the market price of the product along with the time and effort costs of purchasing the product. Likewise, in recreation activities, the amount of participation of individual recreators is determined by the price of the activity and the time and distance costs associated. Again the utility of the experience must be equal to the maximum cost recreators are willing to incur.

In the construction of demand estimates for outdoor recreation that approximate those of market goods, the determination of money-cost prices is essential. Once these costs are defined, demand can be estimated statistically in the same manner as estimations for market commodities. Conventional demand schedules relate alternative quantities of a commodity that would be purchased at alternative prices at a given point in time.

The basic principle of demand states that quantity and price vary inversely. When price is low, large quantities are demanded while when prices are high, smaller quantities are demanded. A typical linear demand curve would appear as shown in Figure 2.1.

A change in price of the commodity causes a change in the quantity demanded of the commodity, all other things held constant. Shifts in demand occur when factors other than the price of the commodity are altered. Such factors include changes in income, population, prices of other competing or complementary commodities, consumer tastes and preferences, changes in time and effort costs associated with purchase, and changes in quality of the commodity.

## Recreation Demand

Recreation demand is simply a modification of conventional demand. The relationship between price and quantity remains unchanged. For recreation demand, the quantity variable is usually in terms of a measure of the use of the resource, for instance visits, trips or user-days. The major difference between demand for recreation and ordinary market-priced consumer demand is that of defining prices. In the conventional type of demand, price is determined at the market equilibrium point where supply equals demand. Most forms of outdoor recreation have no market mechanism for price determination. Consumer prices are absent or set by administrators.

Wennergren (1967) showed that although recreation developed as a nonmarket good, it is not a free good. Time and money costs associated with consumption of recreation activities regulate the quantity of

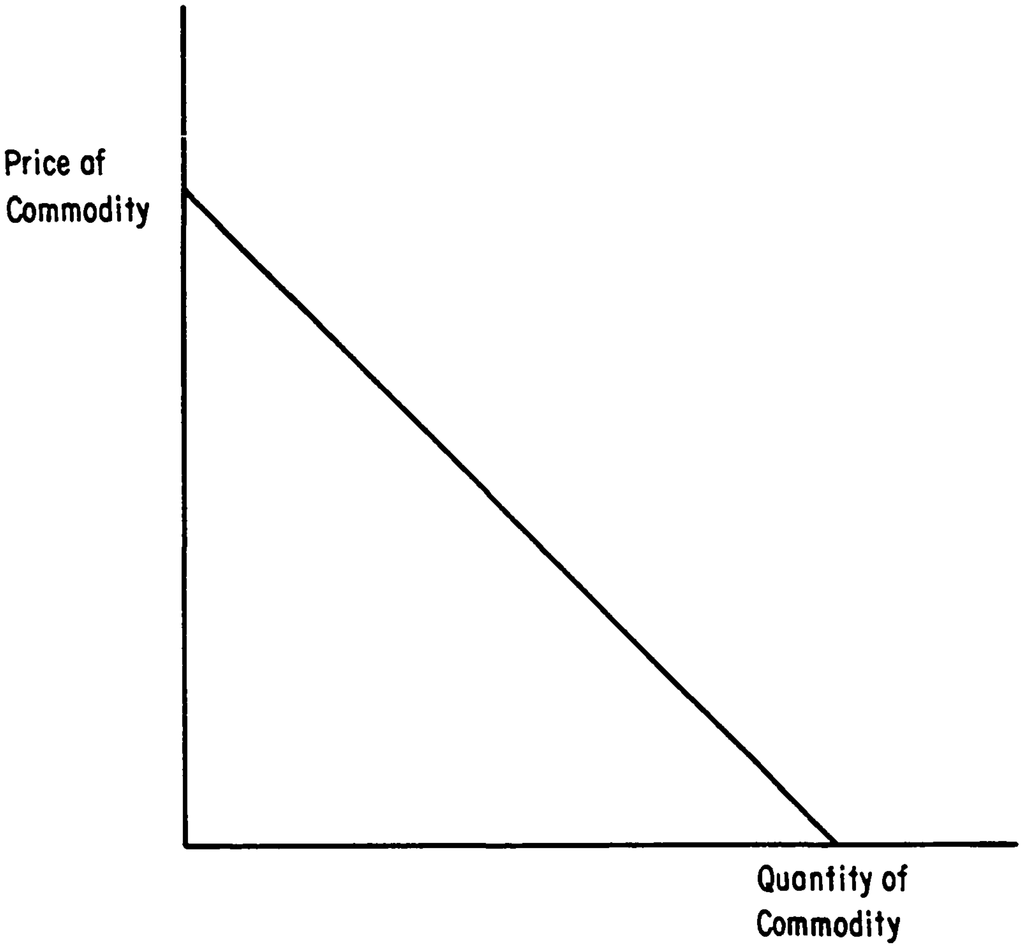


Figure 2.1. Typical linear demand function.

activites consumed. These money costs can be used as surrogate prices in determining demand functions. The time costs associated with consumption can be used as demand shifters.

In determining which costs are to be included as a surrogate price certain distinctions must be made. The potential recreator faces both long and short run cost-related decisions. The long run decisions requires the purchase of items of a fixed nature such as a recreation vehicle and camping equipment, which may be used on more than one outing. Once these fixed costs are incurred they are not affected by short run decisions to participate in a recreation activity. In turn, once fixed costs are incurred they are normally assumed not to effect the short-run decision to demand a recreation activity.

In the short run, the potential recreator faces the decision of which activity in which to participate and at which site. Travel time, travel cost and any on-site costs incurred in the chosen activity are decision criteria. Travel costs and additional variable expenditures are the variable costs and are the pertinent costs for the surrogate price.

Analogous to the use of variable costs in surrogate pricing is the use of variable costs in short-run decision making of the business firm. In the short run, the firm examines marginal costs (additional costs) in making the decision of how much to produce. Marginal costs are a function only of variable costs. Fixed costs do not enter into the production decision. Similarly only variable costs in the short run are pertinent in recreation demand estimation.

## Estimating Demand for Outdoor Recreation

Studies dealing with the availabilities and economic values of outdoor recreation date back to the late 1940s. Beginning with the Prewitt Report (1949) on recreation in national parks, an increasing interest and volume of effort has been recorded in this area. The travel cost method (TCM) of valuing the recreation experience was first suggested by Hotelling in 1949. The initial practical framework for estimating recreation demand was provided by Clawson (1959). In the early 1960s extensive estimates of recreational activity and pragmatic analyses were undertaken by ORRRC (Outdoor Recreation Resource Review Commission, 1962). ORRRC directly addressed problems of demand estimation in several reports. From 1962 to 1967 researchers in many Western states became involved in estimating the economic demand and value for outdoor recreation with emphasis on methodological problems associated with the Claswon-Hotelling approach to demand studies. In 1967, the first regional research project (WM-159) was founded in the Western states to examine outdoor recreation demand.

Two methods of estimating the demand for nonmarket goods including outdoor recreation appear extensively in the literature and are recommended techniques in the Principles and Standards for Water Resources Planning (U.S. Water Resource Council, 1979). These are the travel cost method and the willingness to pay approach.

### Travel Cost Method

Hotelling (1949) proposed that complementary market goods used for the recreation experience could be used to estimate the actual value of the experience. Modifications of this method were suggested by

Clawson (1959, 1966). The travel cost method has become widely known as the Clawson-Hotelling method. The basic elements of the Clawson-Hotelling two-stage procedure for estimating the demand for recreation has been described in detail by Brown, Singh, and Castle (1964), Wennergren (1967a) and more recently, specifically for an Arizona application, by Martin, Gum and Smith (1974).

Clawson (1959, 1966) posits that the use of the recreation site is only one part of the total recreation experience. According to Clawson, the whole experience includes anticipation and preparation for the trip, travel to the site, the on-site experience, and recollection of the experience. Clawson maintains that the recreator gains utility from each of the five phases of the experience. It is the total experience to which the recreator relates the sum total of his expenditure. When measuring the value of an outdoor recreation site, one must separate out the value of the site from the value of the whole experience. The separation involves a two-step procedure. Step one determines the demand for the total experience. Step two derives the demand for the site or activity.

In step one the variable costs associated with the recreation experience are regressed against some per capita quantity variable such as permits or visits to the recreation site from various distance zones. The statistical demand curve resulting from the regression depicts the five phase total recreation experience.

The demand for the recreation resource or site is derived from the statistical demand curve in step two. By making the assumption that recreators will react to an increase in entrance fee for the activity in the same way as they react to an increase in the variable costs associated with travel to the site, conclusions can be drawn about the relationship between entrance fee and quantity demanded. The reaction of recreators to various entrance fees depicts the demand for the site. In developing the demand curve for the resource, the total projected number of visits is calculated at each posited increased interval of cost. The resulting demand curve is in terms of added costs and total quantities of visitation. Observed total visits made to a site from all zones is the point on the derived demand curve of zero added cost. The quantity decrease associated with an increase in added cost can be found for each zone by reading the cost (initial zone costs plus added incremental cost) estimate from the demand curve for the total experience. This cost has a quantity associated with it expressed as visits per capita. Total visits from a distance zone, at this new level of added cost can be calculated as follows:

$$\frac{\text{visits}}{\text{capita}} \times \begin{matrix}\text{population of}\\ \text{distance zone}\end{matrix} = \begin{matrix}\text{total visits from}\\ \text{distance zone}\end{matrix}$$

This calculation is carried out for each distance zone. Total visits are summed over all zones and this quantity is plotted against the added cost increment as one point on the derived demand curve. This procedure is repeated for different added costs to arrive at a derived demand schedule for the resource.

The Clawson-Hotelling approach has been modified and improved over time by many researchers. Ward (1980) suggested that some travel costs are endogenously determined by household tastes and preferences.

He incorporated these socioeconomic variables into a model using a series of simultaneous equations.

Smith and Kopp (1980) proposed that a statistical test for the stability of estimating demand function parameters should be used to define the spatial limits of TCM. The test estimates over what travel range parameters are significant.

Allen, Stevens, and Barret (1981) investigated the effects of omitting variables such as travel time and congestion from the TCM model. They concluded that effects of variable omission may be indeterminable. It is more widely held that omission of these variables may cause TCM to underestimate the true value of recreational resources (Gum and Martin, 1965) (Wetzel, 1977) (McConnell and Duff 1976).

Several recreation demand studies have provided direct application of modified Clawson-Hotelling models to the hunting and fishing industries. One of the pioneering efforts in the field is the study of the Oregon salmon and steelhead sport fishery by Brown, Singh, and Castle (1964). In 1972, Brown and Nawas showed how estimates of value could be improved significantly over the traditional approach by using observations on individual recreators rather than averaging individual observations within distance zones. Martin, Gum, and Smith (1974) confirmed these findings by using continuous distance parameters obtained from individual travel observations to circumvent the behavior-aggregating effects associated with the use of distance zones.

In 1983, Brown, Shorhus, Chou-Yang, and Richards reconsidered the Brown and Nawas (1972) article on individual observations and offered two cautionary suggestions. They suggested that the use of individual observations can lead to biased demand and consumers' surplus estimates unless observations are adjusted on a per capita basis. The use of nonadjusted individual observations implicitly implies that the percentages of participation from each distance zone are equal. The smaller percentages of the population who participate from more distant zones are not properly accounted for. To avoid bias it is suggested that each individual observation be divided by its zone's proportion of the population and be expressed on a per capita basis. A second caution forwarded in the study concerns measurement errors in the variable cost of travel. It was found that biases in the demand estimates due to measurement errors were substantially lower when distance zone average observations were used as opposed to either adjusted or unadjusted individual observations. Thus, the use of zone averaged observations yielded more accurate consumers' surplus estimates.

In 1974, Martin, Gum, and Smith studied all types of rural outdoor recreation activities in all regions of Arizona. This focus made possible the inclusion of prices of substitute recreation attractions as variables in the estimated demand equations. An important emphasis of the report (Martin, Gum, and Smith, 1974) was on demand for and value of a recreation area and activity to a household. The authors acknowledged the household as the prime decision making unit in choosing recreation activities. It was posited that if only the expenditures of licensed sportsmen were recorded, underestimation of actual expenditures to hunting and fishing would occur. In 1975, Martin and Gum summarized the state of the art of TCM. In 1981, TCM was compared to the willingness to pay and the unit day value approaches of demand estimation with respect to competition between cattle and elk for rangeland in Arizona (Heifrich, 1981).

Capal and Pandry (1972) used an additional quality variable in the TCM model. Quality was defined as the hunter's success ratio: the ratio of the number of animals killed to the number of visits to a site.

## Willingness to Pay Approach

The willingness to pay approach (WTP) to recreation demand estimation is designed to estimate benefits from project development due to changes in recreation opportunities or quantities demanded. WTP is also known as the Contingent Valuation Method (CVM). Through a variety of surveying techniques, the researchers ask people directly what they are willing to pay for varying amounts of recreation resources. Randall (1977) led the development of WTP approach for determining the Marshallian demand curve. He advocated the use of bidding games in estimating WTP. Bidding may be an iterative process, such as in interviewing, or noniterative, such as in mail surveys (Randall, 1977) (U.S. WRC, 1979).

The assumption is made that when a person is willing to pay a certain amount, he is also willing to pay less than that amount for the same quantity. For instance, if a person bids $3.00 for a good he is counted amoung the quantity of people willing to pay $3.00 or less. In this way WTP creates a cumulative function relating quantity of respondents to price that is the demand curve for the recreation site or activity. The distinct advantage of WTP method is that the demand curve for the site can be derived directly through the analysis.

There are inherent disadvantages associated with the surveying techniques of WTP. Inaccurate responses or biased responses due to the form of the survey or the perceived results of the survey must be considered. For instance, if an interviewee expects the survey results to increase the quantity or quality of a recreation experience without affecting his entrance fee, he may tend to adjust his true value upward. Likewise, if he expects that the outcome of the survey will effect the cost he pays for a recreation experience he may tend to adjust his true value downward. Biased and inaccurate responses can be minimized when surveys are carefully constructed to explain the purpose of study and the outcome of the results and when questions are phrased to verify information aquired from the interviewee in previous responses.

## Recreation Resource Value

There are two generally recognized methods for estimating recreation resource value. The first is to approximate consumers' surplus implied by an estimated recreation demand function. The alternative is to use the nondiscriminating monopolist value.

### Consumers' Surplus Value

Consumers surplus measures the surplus satisfaction that a consumer receives from a commodity above the price that he actually pays for that commodity. The idea behind consumers' surplus is that a consumer has in mind a price he is willing to pay for a commodity. The price he has in mind must be greater than or equal to the actual price he pays or the consumer would choose to go without the commodity. Since price is a measure of utility or satisfaction, the difference between the price the consumer is willing to pay and the price he actually pays is a measure of surplus satisfaction.

In Figure 2.2 consumers' surplus is represented by the shaded portion of the graph. Consumers' surplus is the area under the demand curve that is above the market price of the commodity (Po). Examining the graph, at least one person is willing to pay a higher price of $P^*$ for the commodity. This individual's surplus benefit at the quantity he consumes is $P^* - Po$. Other persons are willing to purchase the commodity at lesser prices than $P^*$ until finally at a zero price all persons ($Q^*$) would purchase the commodity. Mathematically, consumers' surplus is determined by evaluating the integral under the demand curve up to the quantity (Qo) that is being sold, and subtracting out the total cost (PoQo) that is actually being paid.

Since there is essentially no entrance fee for most outdoor recreation activities, the entire area under the demand curve for the site measures the quantity of consumers' value. This demand curve for a recreation site and its consumers' surplus are depicted in Figure 2.3. In a deer hunting study, Wennergren (1967) defines this area as the quantitative estimate of the value of the marginal utility generated by the resource. This is true because recreationalists equate the value of the marginal utility gained in participation with the marginal cost of participating. Since all visits except the marginal visit were worth more in terms of quality than associated costs, the surplus satisfaction or consumers' surplus is a measure of the net economic value generated from the resource.

Wennergren (1967b) points out that consumers' surplus is the worth of the recreation resource extracted by recreators over and above the cost of participation in the experience. Since this value is not generally involved in market mechanisms, and does not influence the region's economic activity, the worth of the resource cannot be thought to arise from or be expressed in monetary prices in the traditional sense. Consumers' surplus can, however, be expressed in monetary units as long as it is recognized that these units measure some amount of satisfaction above cost (Wennergren 1967b).

The concept of consumer's surplus was first discussed by Depuit in 1844. He describes surplus in monetary terms as the difference between the amount of money a consumer would willingly forgo and the amount of money he actually forgos in order to purchase a commodity (Currie, Murphy, and Schmitz, 1971, p. 742). Marshall (1936) is most often accredited with the development of the concept of consumer's surplus because he defines it in terms of utility rather than prices (Currie et al., 1971). He extended the concept to an aggregate concept which made possible the study of welfare changes among groups.

Hicks (1943) furthered the notion of consumer's surplus by examining the effects that changes in consumer's opportunity sets have on changes in the consumer's welfare. Changes in opportunity sets can be brought about by policy decisions to redistribute goods or income or by price changes. Hicks defined four alternative measures of consumer welfare. These are equivalent surplus (ES), equivalent variation (EV), compensating surplus (CS), and compensating variation (CV). All four

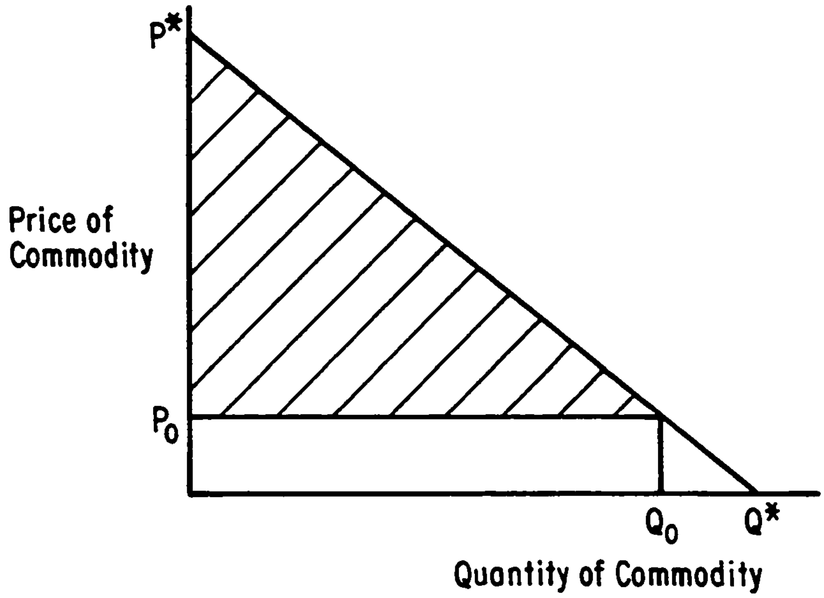


Figure 2.2. Relationship of consumers' surplus.

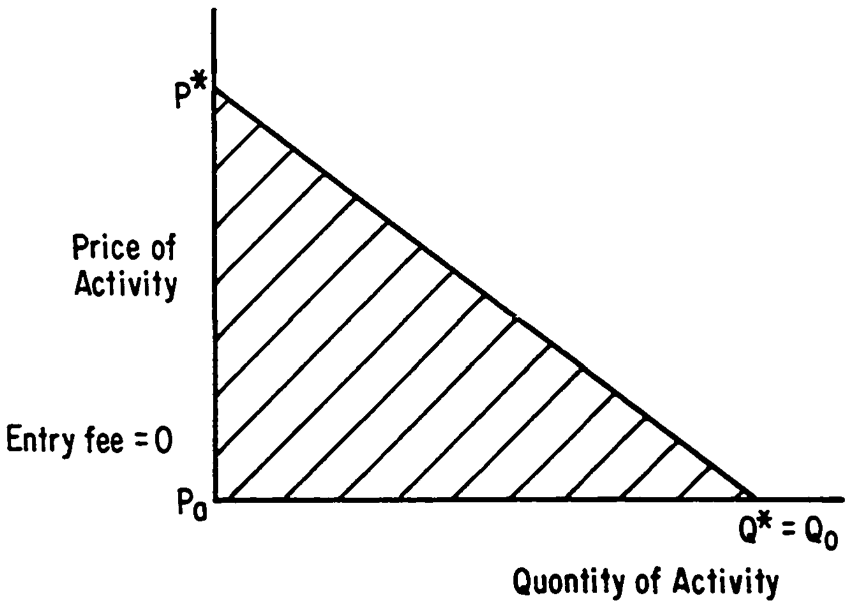


Figure 2.3. Consumers' surplus in outdoor recreation.

measures account the amount of compensation necessary to bring a consumer to some welfare level given various circumstances of price and commodity availability.

ES is defined as the amount of compensation necessary to bring the consumer to his subsequent welfare level in the absence of a price change if he is not permited to make changes in his bundle of goods.

EV is defined as the amount of compensation necessary to bring the consumer to his subsequent welfare level in the absence of a price change if he is permitted to make changes in his bundle of goods.

CS is the amount of compensation necessary to leave the consumer in his original welfare state after a price change if he is not permitted to make changes in his bundle of goods.

CV is the amount of compensation necessary to leave the consumer in his original welfare state after a price change if he is permitted to make changes in his bundle of goods.

The difference between the equivalent measures and the compensating measures is a matter of assignment of rights to the consumer to his initial or subsequent welfare states given some change has occured that affects his welfare. Equivalent measures assume that the consumer has rights to his subsequent welfare position or level of utility. Compensating measures assume the consumer has rights to his original level of welfare or utility.

The relationships among the measures described depend on whether welfare changes arise from a price increase or a price decrease. For a price increase in a normal good ES   EV   Marshallian consumers' surplus CV   CS. For a price decrease ES   EV   Marshallian consumers' surplus

CV   CS. When no income effect arises from a price change, or no change in welfare arises from a change in consumer's opportunity sets, all five measurers are equivalent.

Willig (1976) has shown that CV, EV, and Marshallian consumers' surplus are insignificantly different when the proportion of a consumer's total income spent on a good is small or when the income effect from a change in the good's price is small.

Since the process of estimating simple Marshallian consumers' surplus is less complicated than the estimation of CV, and since both methods yield comparable results, simple consumers' surplus provides a more practical estimate of resource values associated with hunting demand.

Consumers' surplus values the maximum net value of the resource site to consumers when used for a particular purpose such as hunting. Resource planners usually need to compare values of alternative activities occuring on a single site. Often consumers' surplus estimates for alternative uses of a resource are not available with which to make comparison. For purposes of comparison it is sometimes more useful to be able to assign a single price to a value estimate. In these cases an alternative valuation method is used (Martin, Gum, and Smith, 1974).

Nondiscriminating Monopolist Value

The nondiscriminating monopolist method allows for the introduction of a single price that reflects resource value. The model assumes that there is a single monopolistic owner of a resource who wishes to charge the price that maximizes his total revenue from the resource. Since the monopolist charges a single price he cannot discriminate among consumers by extracting from each their total willingness to pay. If the monopolist could discriminate totally, he would extract the entire consumers' surplus.

The price where total revenue is maximized corresponds to a point on the demand curve where the elasticity of demand equals unity. Figure 2.4 shows the relationship between total revenue and demand. The price and quantity of the commodity offered by the monopolist are $P_o$ and $Q_o$, respectively. The shaded area represents the total revenue extracted by the monopolist. Brown, Singh, and Castle (1964) describe this quantity as the net economic value of the resource accruing to some single owner. When the single owner is a resource planning agency, the value of the resource is the maximum net revenue that could be generated by charging a given fee. The fee need not actually be charged for the resource value to be estimated. In fact, the implementation of such a fee might raise serious equity considerations.

Martin, Gum, and Smith (1974) used this method to estimate values for all types of outdoor recreation in Arizona. They found that the monopolist prices for 1970 varied from $5.00 per household-trip for general hunting to $60.00 per household-trip for deer hunting. Nondiscriminating monopolist values were found to be about one-fourth to one-half of the corresponding values estimated by consumers' surplus method.

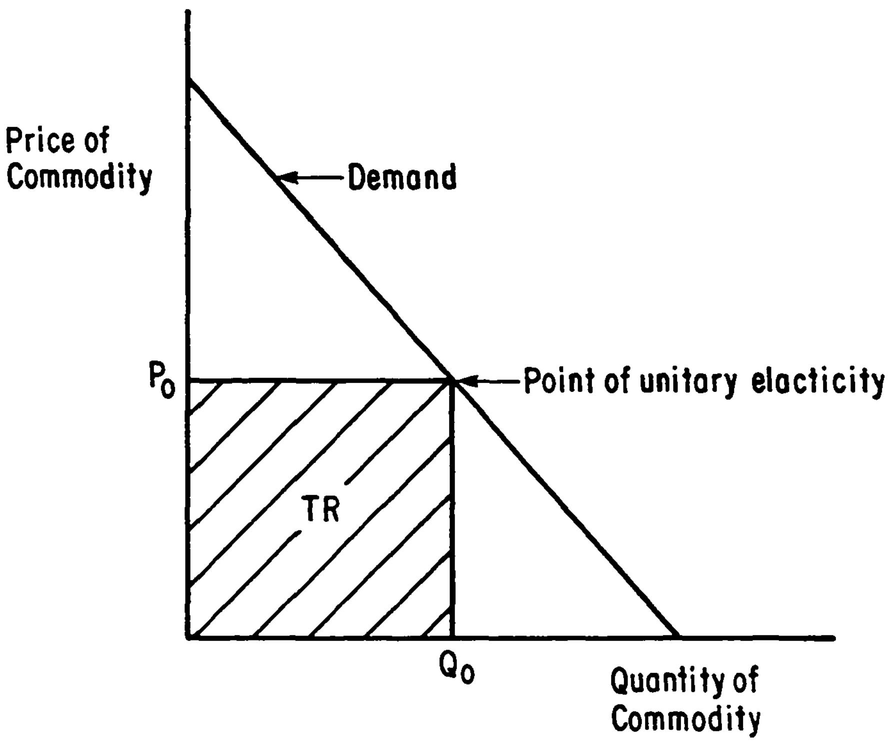


Figure 2.4. Relationship between demand and total revenue.

CHAPTER 3

THE LOG-LINEAR MODEL

## Cross-Classified Categorical Data

Log-linear models are used to analyze cross-classified categorical data. The data are cross-classified under several characteristics into tables of counts called contingency tables. Each axis of the table represents a single characteristic or variable. Each variable is subdivided into ranges or set of categories. For instance, the variable age might be divided into the categories young, middle-aged or old. Data categorized in two or multidimensional tables creates special analytical and interpretive problems.

Bartlett (1935) first addressed the problems of analyzing data compiled in 2 x 2 x 2 contingency tables. But until recently the analysis of multi-way tables has been limited. Killion and Zahn (1976) have compiled an extensive bibliography on the statistical literature pertinent to contingency table analysis through 1974.

The analysis of cross-classified categorical data falls within the broader framework of multivariate analysis. Regression analysis, and analyses of varience and covarience are all forms of multivariate analysis. These latter analyses make a distinction between response variables and explanatory variables. Explanatory varibles are independent or fixed by experimental design because they play a causal role in determining the situation under study. Response variables are dependent and free to vary in response to conditions set by the explanatory variables. Yet, Depster (1971) points out that the relationship between response and explanatory variables need not be fixed in a given situation or specified by design. In analysis of cross-classified data the response variable need not be explicitly stated.

In the method of examining interactions among the data, log-linear analysis is similar to analysis-of-varience (ANOVA) models. But Feinberg (1981) points out that such analogies are deceptive. There is an important distinction between contingency table analysis and other forms of multivariate analysis which often goes unnoticed. ANOVA models seek to assess the effects of independent variables on dependent variables. Contingency table analysis seeks to describe the structural relationship among the (independent) variables corresponding to the the table's dimensions.

This property of defining the structural relationship among variables in contingency table analysis makes log-linear modeling a desirable tool in economic demand estimation. Often demand curves are estimated using regression analysis. When the number of independent variables could be large, specification of the most appropriate model can be a hit-or-miss proposition. Log-linear analysis has the advantage of separating out significant variables and interactions for examination prior to the specification of a model. The technique has the additional advantage of indicating the general shape of the demand curve prior to the specification of the mathematical function.

## Log-Linear Analysis

The purpose of the log-linear analysis is to obtain a description of the relationships among the variables in the contingency table.

The log-linear model is fitted to cell frequencies within the table. The logarithm of the expected cell frequencies then is written as an additive function. An appropriate model for the data can be fitted either by prespecifying model types or by using a screening effect.

Central to the analysis is the test for independence between rows and columns (or variables) of the table. In a two-way table the test of independence is synonymous with the test of fit of a model. Starting with the hypothesis that rows and columns are dependent, statistical tests are performed to test the hypothesis. A significant test statistic implies that the variables tested are highly dependent, and therefore, the model of independence fits the data poorly. Test statistics which are nonsignificant imply good fitting models. Under the hypothesis that rows and columns are independent, the expected frequency in each cell is calculated as

$$E(f_{ij}) = \frac{(\text{ row total}) \times (\text{column total})}{\text{total of table}}$$

For simplicity we can express the expected cell frequency as the product of two terms, one which depends only on the row ($\alpha_i$) in which the frequency appears and the other ($\beta_j$) which depends only on the column. In higher-way tables the expected cell frequencies may be expressed as the product of several terms. For a three-way contingency table with dimensions I, J, and K the expected cell frequency can be written as

$$E(f_{ijk}) = \alpha_i \; \beta_j \; \gamma_k$$

Taking the logarithum of both sides of the above equation yields the general log-linear model. Since the log of a product of terms is the sum of the logarithms of the terms, the log of the expected frequencies can be expressed in the following linear model;

$$l_n\ E(f_{ijk}) = \Theta + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC} + \lambda_{ijk}^{ABC}$$

and where the following constraints hold true:

$$\Sigma_i \ \lambda_i^A = 0 \qquad \Sigma_i \ \lambda_{ij}^{AB} = \Sigma_j \ \lambda_{ij}^{AB} = 0$$

$$\Sigma_j \ \lambda_j^B = 0 \qquad \Sigma_i \ \lambda_{ik}^{AC} = \Sigma_k \ \lambda_{ik}^{AC} = 0$$

$$\Sigma_k \ \lambda_k^C = 0 \qquad \Sigma_j \ \lambda_{jk}^{BC} = \Sigma_k \ \lambda_{jk}^{BC} = 0$$

$$\Sigma_i \ \lambda_{ijk}^{ABC} = \Sigma_j \ \lambda_{ijk}^{ABC} = \Sigma_k \ \lambda_{ijk}^{ABC} = 0$$

The $\lambda$'s are the effects associated with the variables in superscript. For instance, $\lambda^A$ is the effect due to variable A alone. This is termed a first order effect because only one variable is involved. Similarly $\lambda^{ABC}$, the effect of the interaction of variables A, B and C, is a third order effect.

Since all possible effects of and between variables are contained in the model specified above, it is said to be saturated. By setting certain effects equal to zero, different models can be formed. The log-linear analysis used in this study considers only heirarchical models. In heirarchical models, higher order effects such as $\lambda^{ABC}$

cannot be included unless all lower order effects such as $\lambda^{AB}$, $\lambda^{AC}$, $\lambda^{BC}$, $\lambda^{A}$, $\lambda^{B}$, $\lambda^{C}$ and $\theta$ are also included in the model. Including a higher order effect implies the inclusion of all lower order effects. In this way, models can be described by a minimum number of effects (BMDP, 1981).

In understanding the relationship among the variables, the analysis is undertaken in three distinct stages:

1. screening for an appropriate model,
2. testing and comparing models under consideration, and
3. examining disparities between observed and expected values under a chosen model.

When the contingency table is two or three dimensions, the easiest way to screen for an appropriate model is to examine goodness-of-fit statistics associated with all possible models. For a three-way table, 17 relevant models exist. They are as follows:

| | | |
|---|---|---|
| [A] | [AB] | [AB, AC] |
| [B] | [AC] | [AC, BC] |
| [C] | [BC] | [BC, AB] |
| [A, B] | [A, BC] | [AB, AC, BC] |
| [B, C] | [B, AC] | |
| [A, C] | [C, AB] | |
| [A, B, C] | | |

A final model [ABC] exists but is not included here because it defines a nonspecific relationship among the variables. The saturated three-factor effect relates all three variables in such a way that the interaction between any two variables depends on the value of the third variable. In such a model the relationships among the variables cannot be stated in a straightforward manner. There exists an effect associated with each level of a variable and thus each cell of the tube.

Once the expected cell frequencies have been estimated for all models, the goodness-of-fit statistics for the respective models can be examined by either the Pearson goodness-of-fit chi-square statistics ($X^2$) or the likelihood ratio statistic ($G^2$).

The Pearson chi-square is calculated

$$X^2 = \sum \left[ \frac{(\text{observed-expected})^2}{\text{expected}} \right]$$

where summation occurs over all the frequencies in all cells of the contingency table. A large value of $X^2$ corresponds to a value in the right-hand tail of the chi-square distribution and is indicative of a poor fit. Fienberg (1981) warns that the practice of comparing the Pearson chi-square statistic to the tail values of the chi-square distribution at the appropriate degrees of freedom is an approximation that is appropriate only when sample size N is large.

Feinberg (1981) defines large as at least ten times the number of cells in the frequency table. Yates (1934), Cox (1970), and Rao (1973, p. 414) describe continuity corrections for use in 2 x 2 contingency tables. Since continuity corrections for multidimensional tables are at best complicated, none are considered for this analysis (Fienberg, 1981, p. 22).

The second estimate of goodness of fit of a model is the likelihood ratio statistic ($G^2$). This estimate is used to compare models when one model is a special case of another. Suppose that the expected values for observed frequency $x_{ijk}$ are

$$E_{ijk}(1) = (\text{Expected})_1$$

$$E_{ijk}(2) = (\text{Expected})_2$$

and model 2 = [A][BC] is a special case of model 1 = [A][B][C].

The likelihood ratio would be estimated

$$G^2 = 2\sum (\text{observed}) \log \left[\frac{(\text{Expected})_1}{(\text{Expected})_2}\right]$$

The ratio tests whether differences in the expected values are due to random variation if Model 1 is the correct model. Degrees of freedom for this ratio are calculated as the difference between the degrees of freedom for the two models (Feinberg, 1981, p. 57). A large $G^2$ indicates differences in expected values are due to random variations.

Both the Pearson chi-square statistic and the likelihood ratio sum over all cells in the table. Both are asymptotically distributed as chi-square. Their degrees of freedom can be calcualted as n - p, the number of cells in the table (n) minus the number of independent parameters filled (p), when sample sizes are large and when the model fitted is the most appropriate representation of the data (Feinberg, 1981, p. 41).

If contingency tables are greater than three dimensions, specification and testing of all models becomes exhaustive. Tests of marginal and partial association may be desired to screen effects. The partial association test is to see if the association between a set of factors in a given $\lambda$ effect is equal to zero. For example, to see if the partial association of A with B is zero, the difference is taken between the fit of model [AB, AC, AD, BC, BD, CD] and the model [AC, AD, BC, BD, CD]. Note, $\lambda^{AB}$ has been excluded from the second model. Since this test is obtained by taking the difference between two nested models, only the liklihood ratio ($G^2$) is appropriate (BMDP, 1981, p. 178).

Marginal association tests whether or not the marginal association between a set of factors in a given $\lambda$ effect is zero. In order to perform this test a marginal subtable must first be generated. Marginal subtables are summed over all remaining variables except the two of interest. The test of marginal association tests that $\lambda^{AB} = 0$. By fitting the model (A,B) to the resultant two-dimensional marginal subtable, marginal associations can be screened (BMDP, 1981, p. 180).

The tests of marginal and partial association are equal when examining main effects and highest-order effects. Both tests can be used simultaneously to indicate if various interactions are significant. Interactions would be highly necessary in the model when both tests were highly significant and questionably necessary when only one test proves significant (BMDP, 1981, p. 180).

Along with examination of the goodness-of-fit statistics provided for all models, Benedetti and Brown (1978) suggest another method for choosing an appropriate model. They describe model building in a stepwise manner. Starting from an "overspecified" model, that is, one with an abundance of possible effects, they suggest systematically deleting terms. The log-linear program (BMDP, P4F, 1981) calculates the test-of-fit and test-of-significance between the "new" model and the original model. After fitting all possible new models, the best model can be chosen. The best model is the one for which the test-of-significance for the difference between the two models is the least significant, and thus has the largest tail probability (BMDP, 1981).

Once the best model has been chosen, it is often useful to understand why other models exhibited a significant lack-of-fit. Differences between observed and expected values can be examined directly or in terms of the deviates based on these differences. Examination of the estimates of the parameters can provide insight into relationships among the variables. A number of extreme values may be identified and replaced with structural zeros. The model may then be retested for significance (BMDP, 1981, p. 201).

## Log-Linear Modeling of the Demand for Hunting

A three-way contingency table is constructed using secondary data from individual hunter permit applications for deer and javelina, respectively. Each table has as its axes the decision variables, the probability of draw ($P_d$), the probability of success of kill ($P_s$), and the miles to be traveled (M) from the hunter's residence to the hunting site. Each variable is divided into reasonable categories over which the hunter may discriminate with his behavior or preference. Hunters for a particular animal are classified according to their characteristics (or categories of decision variables) and counted in the appropriate cell of the table. The contingency table then provides the observed cell frequencies of all hunters for a particular animal. A hypothetical contingency table is constructed in Table 3.1. The example contains 16 observed cell frequencies.

In Table 3.1 there are four separate arrays of data indicated by four partitioned areas. Each area in the table has associated with it a specific level of $P_d$ and $P_s$ over a spectrum of miles categories. The observations in each box denote the number of deer hunters applying for a permit to hunt in a particular hunting area that falls within each of the miles categories associated with that area. Each hunting area is defined by some combination of $P_d$ and $P_s$. More than one area may be defined by the same combination of $P_d$ and $P_s$ and thus be defined in the same partitioned area of the table.

Under the hypothesis that the decision variables $P_d$ and $P_s$ are indicative of the quality of the site, we may say that observations in a partitioned area of the table represent the quantity of hunting permits demanded for a particular quality of hunting site at various ranges of miles traveled. These observations reflect the psuedo-demand function for a given quality of site. The psuedo-demand function, is a discontinuous step function relating the quantity of applicants, as determined by the odds of receiving benefits ($P_d$ and $P_s$) to the cost incurred in receiving benefit (miles traveled). Figure 3.1 is the observed psuedo-demand function for the quality of site represented by the upper left-hand partition of the contingency table. The psuedo-demand function is representative of all sites which have probabilites of draw and success in the range of 0.0 to 0.5, respectively.

The next step in the analysis is to determine the expected cell frequencies given that an appropriate model has been chosen. The log-linear program (BMDP, P4F, 1981) generates a contingency table of expected cell frequencies once the appropriate model has been specified. Expected cell frequencies indicate the number of hunter applicants in

Table 3.1. A hypothetical observed frequency table for deer hunters.

| Probability of Success | Miles to Area | Probabilities of Draw | | Total Number of Hunters |
|---|---|---|---|---|
| | | 0.00 - 0.50 | 0.51 - 1.00 | |
| 0.00 - 0.50 | 0-100 | 20 | 25 | 45 |
| | 101-200 | 15 | 20 | 35 |
| | 201-300 | 10 | 15 | 20 |
| | over 300 | 0 | 10 | 10 |
| Subtotal | | 45 | 70 | 115 |
| 0.51 - 1.00 | 0-100 | 35 | 50 | 85 |
| | 101-200 | 30 | 45 | 75 |
| | 201-300 | 25 | 40 | 65 |
| | over 300 | 20 | 35 | 55 |
| Subtotal | | 118 | 170 | 280 |
| Grand Total | | 155 | 240 | 395 |

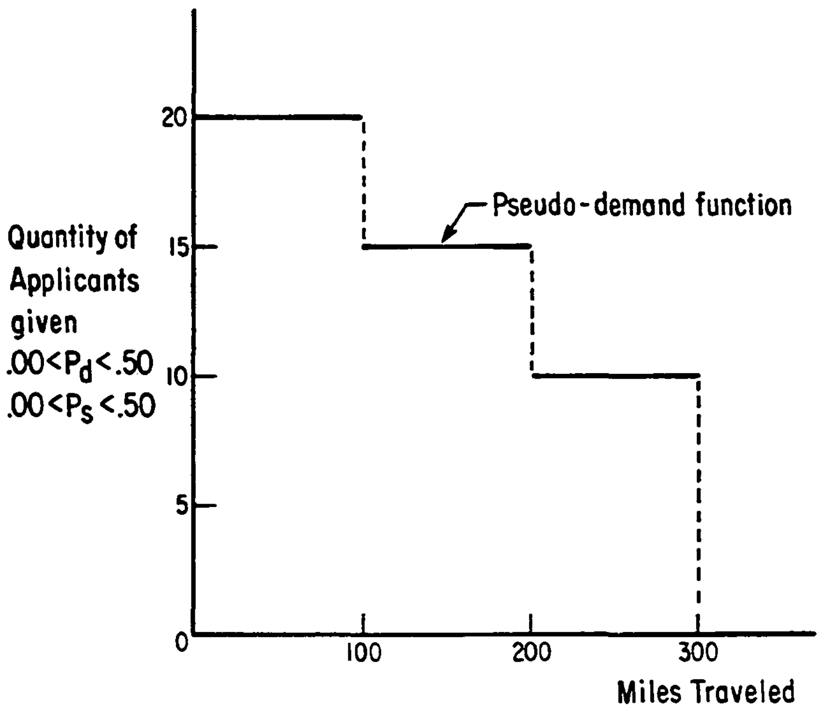


Figure 3.1. A hypothetical pseudo-demand function for a given quality of deer hunting site.

each set of categories predicted by the model. The partitioned areas of the "expected" contingency table are the fitted psuedo-demand curves for the various quality sites.

The relationship expressed by the fitted psuedo-demand curve is a proxy to the cost relationship that is similar to the second stage demand curve estimated by the travel cost method. It is the second stage curve since all people interested in deer or javalina hunting have indicated their willingness to pay (travel distance) given the relative quality of the sites. Only site is chosen, and only one trip is allowed. Since the psuedo-demand curve is a step function it can relate the marginal utility of hunting a particular quality site that is gained by traveling an additional quantity of miles. In the example, the psuedo-demand relates the marginal utility gained from traveling an additional 100 miles in order to achieve a given quality of hunt. Utility can be expressed in terms of any units. Here it would be expressed in terms of miles which are indicative of a monetary variable cost incurred by the hunter. By making the assumption that applicants react to an increase in travel cost in the same way as they would react to an increase in site entrance fee, the psuedo-demand curve relates the decrease in site participation associated with an increased site entry fee.

The valuation of the hunting resource at the site is similar in principle to either the nondicriminating monopolist or the consumers' surplus methods. For the consumers' surplus method the site value is estimated as the area under the psuedo-demand curve. Since the psuedo-demand curve is discontinuous, integration cannot be used to assess the value of the area underlying the curve. Instead the areas underlying each segment of the curve must be calculated and summed over all areas. In the example, the value (in miles) of the site or sites depicted in the psuedo-demand function of Figure 3.1 would be

$$\text{miles} = (20 \times 100) + (15 \times 100) + (10 \times 100) = 4{,}500.$$

This measure of utility expressed in miles easily can be converted to a monetary figure by estimating the variable cost of traveling a mile, including such costs as gasoline, automobile maintenance, and repair. Note that the total utility gained by the 15 hunters who were willing to travel 200 miles as opposed to 100 miles was an additional 15,000 miles of benefits. Each hunter received an extra 100 miles of benefit by incurring an extra 100 miles in cost.

Quality comparisons among sites can be observed as the vertical distance between the psuedo-demand curves. Over a given range of miles, the observed differences in the number of applicants for two sites is an indication of preference among hunters to one quality site as opposed to another. The area between two psuedo-demand curves is an estimate of the increased value (in miles) due to an increase in site quality. Again, the units of the measure of utility gained from site quality are unimportant for making comparisons, but the mileage units could be converted to dollar units under the travel cost hypothesis.

## CHAPTER 4

## DESCRIPTION OF THE DATA

This analysis uses secondary data obtained from the Arizona Game and Fish Department. The department records information about all individuals who apply for hunting permits. On the permit application, individuals indicate their name, address, city and state of residence, zip code, and their first and second choices of areas and days to hunt. As many as four persons can apply for a joint permit to participate in a specific hunt. A hunt is specified by a hunt number which indictes the desired hunting site and time, and the type of animal to be hunted. Since the analysis attempts to examine the total demand for hunting deer or javelina rather than the permitted demand, all resident permit applicants for the two animals are examined. Data on joint permit applications are divided into individual observations. Only the applicants' first choice of hunt is considered to be indicative of the actual demand for an area if hunting areas were not restricted to entry by permits.

Two pieces of information are extracted from each individual's permit application: the individual's zip code and his first choice of hunt. With this information, along with the Arizona Game Survey and Harvest Data Summary published yearly by the Arizona Game and Fish Department, three decision variables can be assigned to each individual. The decision variables, miles traveled to the area; the probability of being drawn for a permit, and the probability of success of kill, are the varibles to be tested by the log-linear method.

## Miles Traveled

The miles traveled variable (M) is calculated as the miles traveled from the applicant's residence to the center of the first choice hunting area. Points of travel departure are defined by the zip code associated with the applicant's place of residence. The 1979 zip code maps used for estimating place of residence were provided by the U.S. Postal Service.

When several zip codes were in the proximity of an area measuring 20 miles in diameter, zip codes were grouped together to form a single point of departure. In sparsely populated areas of the state a single zip code may refer to an area as large as 100 miles in diameter but these observations are not common. The maximum error associated with estimating points of trip departure using zip coded addresses is 50 miles. A more reasonable estimate of the average mileage error is 20 miles.

Points of trip arrival are taken to be the center point of the hunting area indicated by the applicants' first choice of hunt. Each hunt is associated with a hunting unit or combination of units that make up the hunting area. Figure 4.1 shows the location of all hunting areas and the chosen center points. In the case where several units are combined to form a single hunting area, the center point of the hunting

For the hunting season 1980-1981, Table 4.1 shows the combinations of units making up hunting areas for deer and javelina, respectively. Combined units forming a single area appear on the same line of a column.

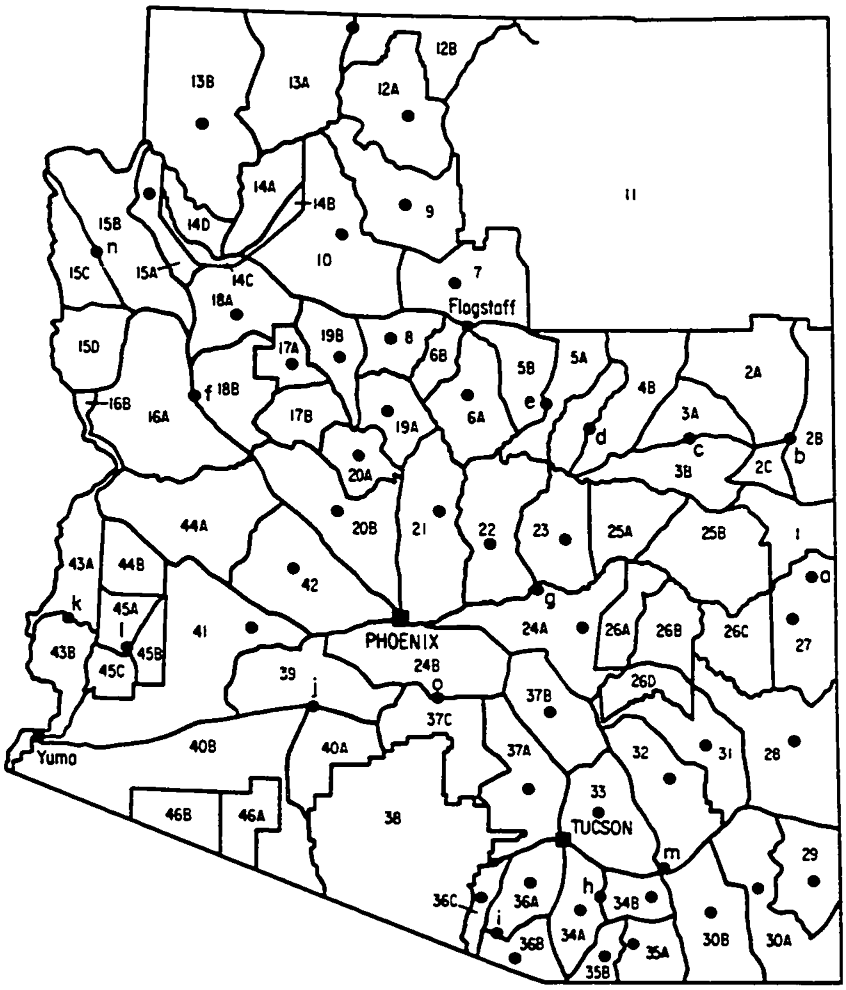


Figure 4.1. Map of center points of hunting areas in Arizona.

Key for center points of combined hunting units.

a. 1, 27
b. 2A, 2B, 2C
c. 3A, 3B
d. 4A, 4B
e. 5A, 5B
f. 16A, 18B
g. 22, 23, 24A
h. 34A, 34B
i. 36B, 36C
j. 39, 40A, 40B
k. 43A, 43B
l. 45A, 45B, 45C
m. 29, 30A, 30B, 31, 32, 33, 34A, 34B, 35A, 35B, 36A, 36B, 36C
n. 15B, 15C, 15D
o. 24B, 37C

The largest hunting area for deer has an approximate radius of 150 miles about its centerpoint. The maximum error associated with using the center of the area as the point of travel arrival is ±150 miles. The average mileage error associated with the average hunting area for deer is approximately ±75 miles.

For javelina hunting areas the maximum error for mileage, calculated from the largest area is ±50 miles. The average javelina area extends approximately 40 miles in any direction from the chosen center point and has an associated average mileage error of ±40 miles.

Mileages were calculated from all potential points of departure to points of arrival using a Arizona Highways 1981 Road Map. Distances were calculated along the shortest or most reasonable travel route. When travel occurred over long distances interstate highways were generally used. A mileage matrix containing 5,320 potential travel routes was created. The data on permit applications were compared with the travel matrix and each individual was assigned a value for the variable "miles traveled."

The maximum error associated with the miles traveled variable for deer hunters would be the maximum error associated with the point of trip departure plus the maximum error associated with the point of arrival. This amount (50 + 150) equals ±200 miles. The average mileage error for deer (20 + 75) is approximately ±95 miles. For javelina the maximum and average mileage errors are ±100 and ±60 miles, respectively.

The miles traveled variable is used as a proxy estimate of the variable costs associated with the hunting area. As is true with the travel cost method of demand estimation, it is hypothesized that the

Table 4.1. Hunting areas available for deer and javelina 1980-1981 hunts.

| DEER | | | JAVELINA | |
|---|---|---|---|---|
| Unit | | | Unit | |
| 1,27 | 18A | 33 | 19A | 36B |
| 2A, 2B, 2C | 19A | 34A, 34B | 20A | 34B, 37C |
| 3A, 3B | 19B | 35A | 21 | 30A |
| 4A, 4B | 20A | 36A | 22 | 34B |
| 5A, 5B | 20B | 36B, 36C | 24A | 36C |
| 6A | 21 | 37A | 27 | 37B |
| 6B | 22 | 37B | 28 | |
| 7 | 22, 23, 24A | 39, 40A, 40B | 29 | |
| 8 | 22 | 41 | 30A | |
| 9 | 24A | 42 | 30B | |
| 10 | 27 | 43A, 43B, 44A, 44B | 31 | |
| 12A | 28 | 45A, 45B, 45C, | 32 | |
| 12B, 13A | 29 | 15A | 33 | |
| 13B | 30A | 15B, 15C, 15D | 34A | |
| 16A, 18B | 30B | 24B, 37C | 34B | |
| 17A | 31 | all units 29-36 | 34A | |
| 17B | 32 | | 35B | |

cost of complementary goods associated with the recreation experience can be used to estimate the value of the experience. The complementary costs associated specifically with a hunting area are predominantly the variable travel costs.

The total variable travel cost to a site is calculated as the number of miles traveled times some constant dollar amount indicative of fuel, maintenance, and repair costs per mile. For purposes of comparing relative travel costs among individuals, it is the same to compare distances traveled by individuals as it is to compare distances traveled multiplied by some constant amount. Thus, the variable travel costs are expressed in this study in terms of miles traveled rather than dollars of cost per trip.

As with the travel cost method, the willingness of individuals to pay varying costs to travel to a site is taken to be the willingness of the individual to assume varying site entrance fees. In the travel cost method, normally a single site is examined. Individuals traveling from various distance zones are associated with different willingnesses to incur travel costs. The total projected number of visits are calculated at each posited increased interval in cost and the demand curve is estimated. In this study, all areas are examined simultaneously. Each individual is associated directly with some variable cost of travel in terms of miles traveled. In a sense, distance zones are related to each individual and centered around each hunter's residence. Each individual chooses the appropriate distance zone he is willing to travel to. Two areas lying equal distances from a hunter's residence would be associated with the same distance zone, cost of travel, and entrance fee. The preference of one area over another, given comparable costs, would be an indication of the individual's perception of a difference in quality between the areas.

## Probabilities

Each hunting area has associated with it two quality variables, the probability of an applicant drawing a permit to the area ($P_d$), and the probability of an applicant killing an animal in the area ($P_s$). It is hypothesized that individuals faced with the decision of choosing among several areas of equal distance from their residence base their decisions on the site attributes relayed by the variables $P_s$ and $P_d$. It is further hypothesized that the hunter uses information about these probabilities that is published in the Arizona Game Survey and Harvest Data Summary (AGF).

These hunting statistics are published annually. Each year applicants have statistics available from the previous season's hunts to use as reference. For example, it is hypothesized that the 1980 deer applicant bases part of his site decision on 1979 season statistics for deer. In this way the previous season's statistics provide the probability data assigned to each applicant for the fall 1980 deer hunts and the spring 1981 javelina hunts.

The data compiled by the Department to complete the Game Surveys (AGF) are obtained from a random survey of hunters who were permitted in the previous season. A questionnaire sent to these hunters includes a letter of explanation and a site location map. For the 1979 deer draw, 41,885 questionnaires were sent and 24,455 or 58.9 percent were returned. Data in the surveys were available for examination before the

1980 hunt applications were due. For the 1980 javelina hunts, 14,909 questionnaires were sent and 56 percent were returned. These data were available before the 1981 application period.

The probability of drawing a permit ($P_d$) for a particular hunt is a function of the number of permits authorized by the department and the number of applicants demanding a hunt. The number of permits authorized depends on the department's determination of the quality of the area, the health and maintenance of the animal population and the desired conjestion in the area. The number of first choice applicants for an area indicates public votes in favor of area quality, location, and possible historic personal experience with the area. The Game and Fish Department publishes for each hunt, the total number of first choice applicants and the number of authorized permits (AGF). The odds of being drawn for a permit can be calculated as

$$P_d = \frac{\text{number of permits authorized}}{\text{number of first choice applicants}}$$

The odds $P_d$ calculated using the previous hunting season's statistics for the hunt of interest must be correlated with the hunts offered in the present season. In some cases a previous hunt on two areas would be divided into two hunts on one area each in the present year. In such cases the best estimate of draw probabilities for this year's two hunts would be to use last year's combined area hunt estimates. Likewise if hunts from a previous year are combined to form a single hunt this year the mean of the two previous hunts can be taken to estimate the present odds of being drawn. It must be assumed that applicants use past statistics to make inferences about their present situation without involving a great deal of mathematics. Permit odds have been calculated and correlated in such a manner and assigned to each applicant.

The probability of a hunter successfully killing an animal on a particular hunt is a function of the magnitude of animal population and the magnitude of the population of hunters present at any time. The number of animals available for kill depends on the natural quality of the area and on the previous season's harvests. The number of hunters in an area is limited by policy decisions made by the Fish and Game Department concerning area conjestion.

The percentage rates of success are published for each hunt by the Department (AGF). This estimate is calcualted as the number of animals killed divided by the number of hunters that visited the area on that particular hunt.

$$P_s = \frac{\text{number of animals killed in an area}}{\text{number of hunters in an area}}$$

In general, hunters stayed in the area between one and two days in order to achieve these success rates. These probabilities calculated for the previous season's hunts must again be correlated with hunts offered in the present season.

In Figures 4.2 and 4.3 each hunt offered for the 1980 fall deer draw is represented by a single dot. Figure 4.2 shows how these hunts are distributed over the range of $P_s$ values. Figure 4.3 shows how these hunts are distributed over the values of $P_d$. Figure 4.4 and 4.5 represent the number of hunts distributed over the probabilities for the 1981 spring javelina hunts.

For the 1980 fall deer draw, $P_s$ estimates range from 0.10 to 0.35. The overall 1979 mean of all deer hunts is $P_s = 0.21$. In 1980, the mean probability of success over all deer hunts was $P_s = 0.22$. For 1981, javelina $P_s$ estimates range from 0.1 to 0.41. Average $P_s$ statistics for all javelina hunts in 1980 and 1981, respectively, are 0.24 and 0.26. It is concluded that using the previous season's statistics to infer present proabilities is on the average a reasonable estimate.

## Applicants

Approximately 87,000 persons applied individually or in groups to hunt deer in the fall of 1980. Some of these applicants were disqualified from the permit draw because the applications were filled out incorrectly. A common error made was in the specification of the desired hunt. Many people indicated the hunting unit desired instead of the hunt number. Since units are used in different time periods for different animals and different "hunts" of the same animal, the applicants' intentions could not be discerned. Of the remaining eligable applicants, 3,691 could not be matched with potential travel routes. A large portion of this number were of nonresident status and are not considered in this analysis. A few Arizona residents could not be matched either because their zip code had been changed or created since 1979 (the year our zip code information was published). The total number of individual observations used for the analysis of 1980 deer hunts is 82,935.

For the spring 1981 javelina hunt, approximately 25,000 individuals made application for permits. Of these, 1,776 persons could not

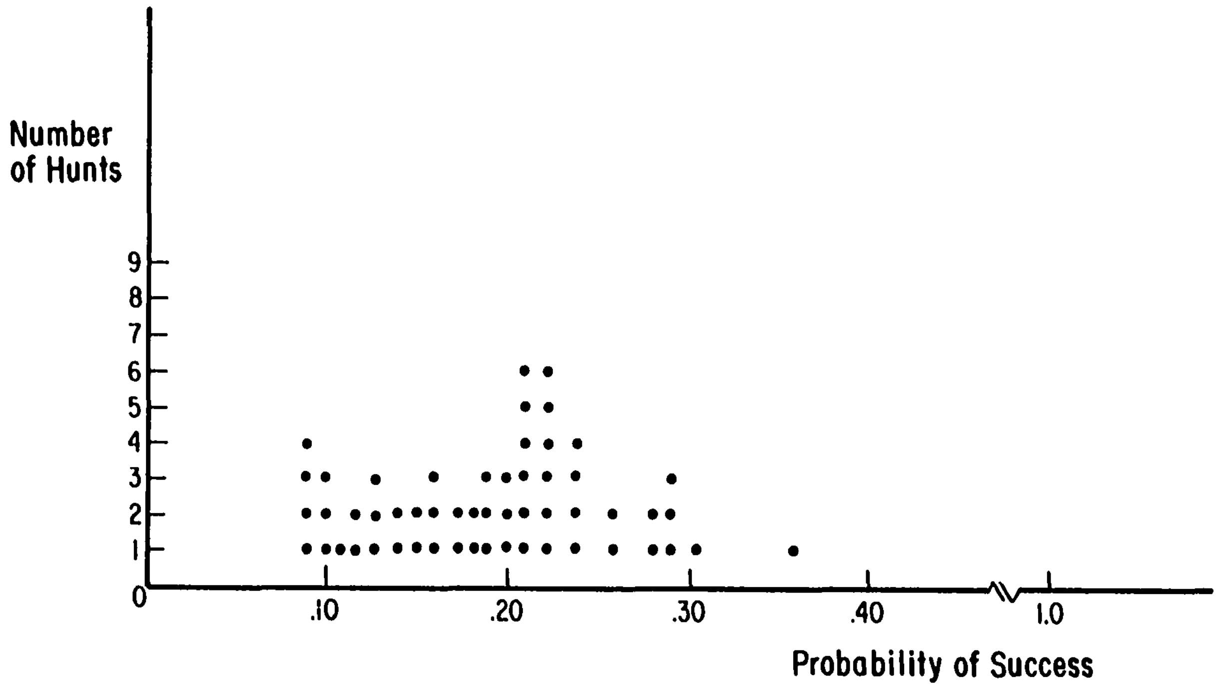


Figure 4.2. Distribution of 1980 deer hunts over the range of probability of success values.

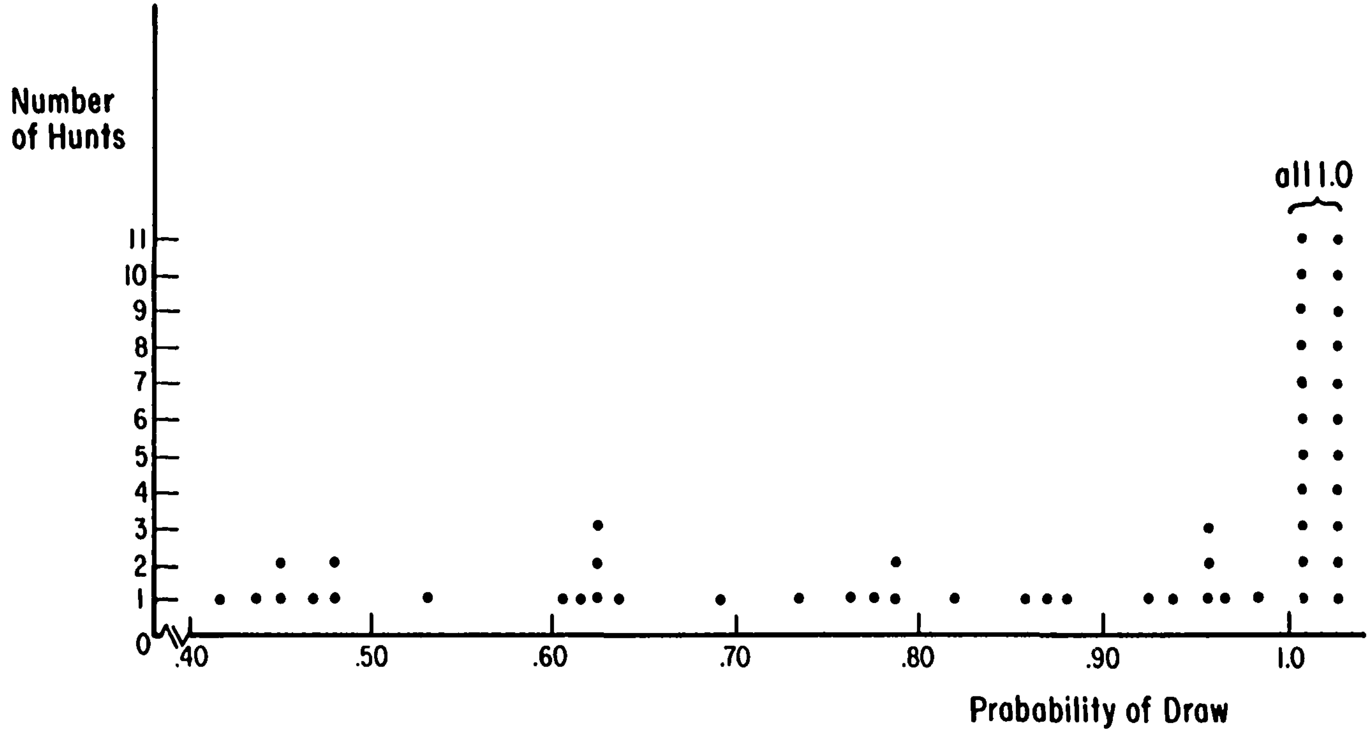


Figure 4.3. Distribution of 1980 deer hunts over the range of probability of draw values.

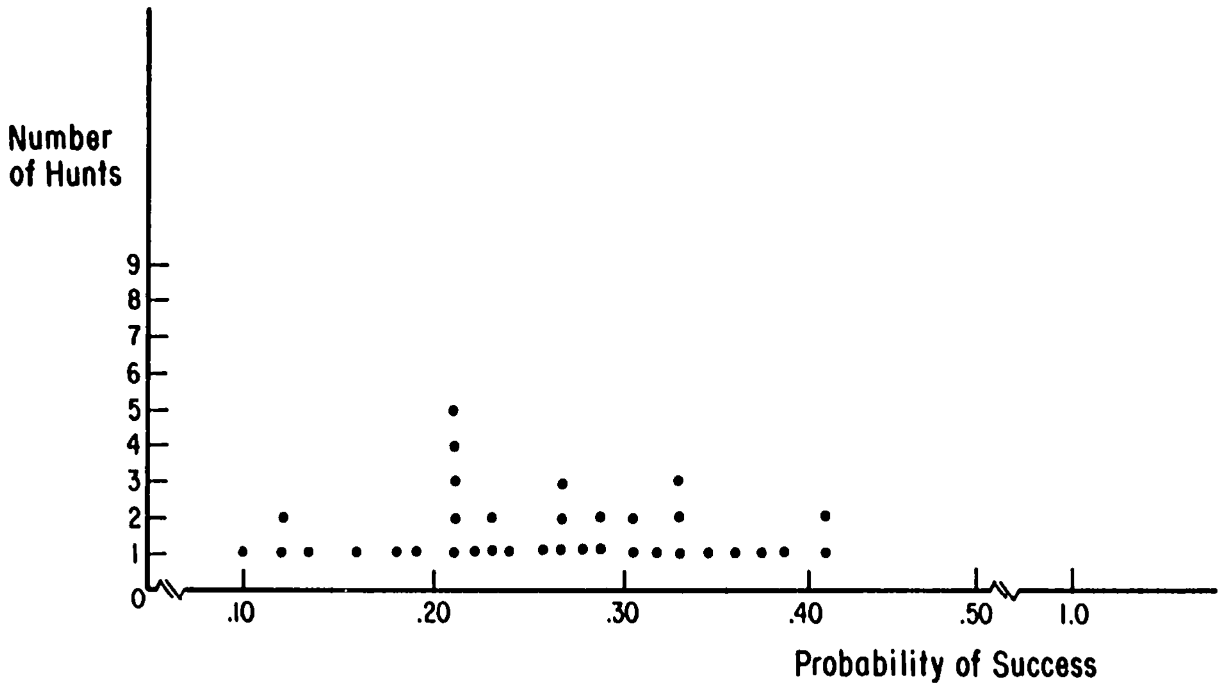


Figure 4.4. Distribution of 1981 javelina hunts over the range of probability of success values.

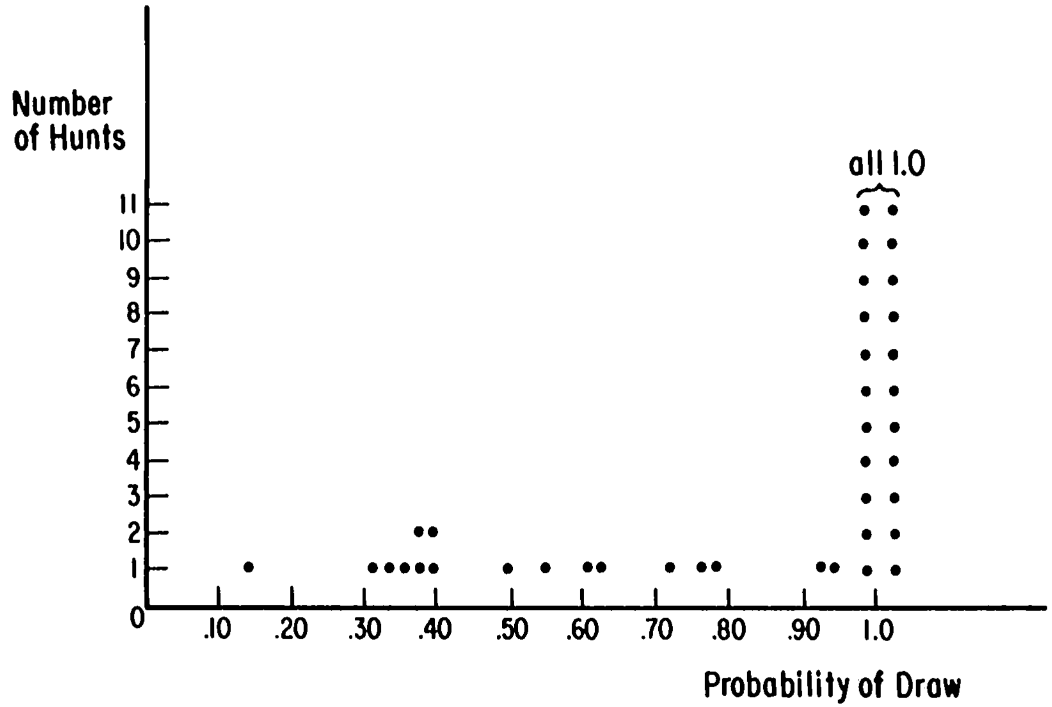


Figure 4.5. Distribution of 1981 javelina hunts over the range of probability of draw values.

be matched with the mileage matrix due to nonresident status or changes in their resident zip codes. After removing these individuals from the total of applicants, as well as those who made application errors, 22,857 individual observations were available for the analysis of javelina hunts.

# CHAPTER 5

# THE RESULTS

## Statewide Model

### Screening the Data for the Appropriate Contingency Table

Prior to the initial stage of the analysis in which the data are screened for significant effects to be included in the model, it is important to define the intervals or categories of the contingency table so that the data are most meaningful. Since one of the primary measures of a model's statistical significance is the chi-square statistic, it is necessary to distribute the data over the intervals in the contingency table in a manner that best approximates a chi-square distribution. Data can be redistributed over the cells in the table by redefining the categories of its axes.

The Biomedical Computer Programs Manual (BMDP, 1981) suggests a rule for inspecting minimum expected values of cell frequencies to determine if variable categories are appropriately defined. No cell should have an expected value less than one and no more than 20 percent of the cells in the table should have an expected value less than five. more cells have smaller expected values then is indicated by the rule, chi-square statistics may differ widely from the chi-square distribution and the estimates of the model's statistical significance lose much of their meaning. Feinberg (1981) suggests a more conservative rule which was initially forwarded by Fisher (1936) and is widely adhered to: No cell may have an expected value less than five. Where the minimum expected values of cells do not meet this criterion, the categories of the table's parameters may be expanded or collapsed until the most appropriate contingency table has been specified.

Another consideration in the construction of contingency table categories which is specific to this analysis is the hunt applicant's ability to differentiate between the categories of any given variable. For instance, applicants may be indifferent to the categories for miles traveled of 50-60 miles and 60-70 miles. When category distinctions are made in the design of the analysis that are meaningless to the applicant information about trends in the data is lost. Information can also be lost when categories are defined too broadly.

Keeping both the statistical and the practical criteria in mind, the categories listed in Table 5.1 were defined for the analysis.

## Screening for an Appropriate Model

Screening for an appropriate model to fit the data can be accomplished by testing all models and all partial and marginal associations of interacting terms. Two contingency tables, one for deer and one for javelina were screened.

Since the contingency tables were only three dimensional, the process of testing all possible combinations of the variables could be undertaken with ease. Seventeen possible models were screened for variable effects. The most simple models test the effects of a single variable alone. Then models of single variables are tested in combinations. The more complex models, second order variable interactions alone and in combinations, are tested last.

Table 5.1. Categories of the decision variables used in the contingency table analysis.

| | -----Deer------ | -----Javelina----- |
|---|---|---|
| Miles (M) | under 25<br>25-50<br>51-100<br>101-200<br>201-300<br>301-400<br>over 400 | under 25<br>25-50<br>51-100<br>101-200<br>201-300<br>301-400<br>over 400 |
| Probabiility of Draw ($P_d$) | under 0.50<br>0.51-0.65<br>0.66-0.80<br>0.81-0.95<br>above 0.95 | under 0.36<br>0.36-0.50<br>0.51-0.65<br>0.65-0.80<br>0.81-0.95<br>above 0.95 |
| Probability of Success ($P_s$) | under 0.10<br>0.1 -0.15<br>0.16-0.20<br>0.21-0.25<br>0.26-0.30<br>above 0.30 | under 0.16<br>0.16-0.20<br>0.21-0.25<br>0.26-0.30<br>0.31-0.35<br>above 0.35 |

For both deer and javelina, the simplest models ([M], [$P_s$] or [$P_d$]) provided the least fit. The first order deer models with approximately 205 degrees of freedom (d.f.), likelihood ratios were approximately 243,000 and Pearson chi-squares were approximately 321,000. These statistics indicate that the variables have highly statistically significant and dependent effects in explaining the data. Since these models are significant it can be concluded that the model of independence between factors does not provide an adequate fit for the data. For these first order models to provide a good fit, the magnitudes of the goodness-of-fit statistics would need to be less than 100. the hypothesis that any respective first order model adequately explains relationships among the data is rejected.

An important result of the first order model tests is that the variable "miles traveled," cannot be used alone to explain applicants' behavior. Viewed on a statewide level, applicants are considering miles traveled to a site and the associated variable travel cost in conjunction with other factors in their decisions to demand certain hunts.

As models of increasing complexity were screened, the test-of-fit statistics were observed to decrease in magnitude, giving the models slightly better fit. However, even in the most complex second order interactions in the models for both animals, models exhibited a distinct lack of fit. These models exhibited a P-value of 0.0 indicating that the model of independence between factors would not provide an adequate fit for the data. This result means that there is complete confidence that the effects relayed by the models are significant and dependent relationships. The results indicate that none of the three decision variables, as hypothesized in the introductory chapter, explain hunt applicants' behavior to demand certain sites in a simple manner. The trend of increasing goodness of fit associated with increasing complexity of the models can be explained as a by-product of the hierarchical method used. Complicated models involving large numbers of parameters most often fit a set of data more closely than simple models which are special cases of the complicated one. Tradeoffs often exist between simplicity in model specification and goodness of models' fit (Feinberg, 1981, p. 56).

According to trends in the levels of signficance among models the best fitting model would be the complex three-way cross product model [$MP_sP_d$]. This model would be associated with nonsignificant Pearsons chi-square and likelihood ratio statistics, a P-value of 1.0 and zero degrees of freedom. This result indicates that while the model of independence between rows and columns provides the best fit in the case of the saturated model [$MP_sP_d$], the effects of the decision variables in explaining the data cannot be ascertained in any simple manner. In the saturated model all factors are related in such a way that the interaction between any two depends on the value of the third variable. For instance, we can say that the interactive effect between the quality variables $P_s$ and $P_d$ depends on whether travel distances are small, intermediate, or large. And that the relationship of either quality variable to miles traveled depends on the value of the other quality variable. This dynamic relationship between factors may typify the game playing strategy of applicants who seek to maximize the perceived benefits derived from some combination of $P_s$ and $P_d$ at some perceived cost in terms of M.

The distinct disadvantage of using the saturated model to understand the relationships among the factors in the contingency table is that these relatinships cannot be stated in a simple mathematical function. The saturated model [$MP_sP_d$] relays as many effects as there are cells in the observed frequency table. Fitting the saturated model would yield an expected contingency table that would be indentical to the observed table. The objective of defining straighforward relationships among factors cannot be accomplished without sacrificing a great deal of the modle's goodness of fit. The next best model that could be chosen in order to define more simple factor relationships would be the model [$MP_s$, $MP_d$, $P_sP_d$]. This model indicates that there is an effect related by the combination of M and $P_s$ that is independent of the effect related by the combination of M and $P_d$, and independent of the effect related by the combination of $P_s$ and $P_d$. Results indicate, however, that the model of independence does not accurately describe the relationships between the data and the model exhibits a distinct lack of fit.

In an attempt to discover nonsignificant interactions among the factors, the tests of marginal association were also performed. Three marginal subtables were created for each set of two variables by summing over the third variable. Marginal association tests determine whether the marginal association of an interaction term such as $\lambda^{AB}$ is zero. By fitting the model [A, B] to the marginal subtable with axes A and B, the test screens interaction among the terms for each category of the third, (omitted) variable. The tests showed that interaction effects among variables were significantly different from zero.

Since all models were tested and proven significant, it was not necessary to perform tests of partial association. Partial association tests to see if the association between a set of factors in a given $\lambda$effect is equal to zero by taking the difference between two nested models which differ only by the $\lambda$ effect of interest. For example, to test that the interaction term $\lambda^{AB} = 0$, the difference is taken between the models [AB, AC, BC] and [AC, BC]. When tests for all models were specified the effect of $\lambda^{AB}$ was tested and proven insignificant. The test of partial association of $\lambda^{AB}$ would be repetitive.

The second stage of the analysis, fitting an appropriate model, could not be performed given that all models exhibited a distinct lack of fit. The final analytical stage involves examining trends in the data which might contribute to a significant lack-of-fit of the tested models. The methods used in this step of the analysis are residual analysis and the partitioning of marginal subtables to examine chi-square statistics. In residual analysis, residuals are calculated as the difference between the observed frequencies and the expected frequencies in each cell of the contingency table. When the absolute value of the residual is large, the value of an observed frequency differs widely from the pattern of observations in the table as a whole. Cells with large residuals can identify data which contribute to a general lack-of-fit of a model. One option in residual analysis is to identify extreme cells and replace the observations with structural zeros. Models then are retested for significance. This option was not open in this study because a large percentage of the data were associated with large residuals and extreme values.

For deer, the largest residuals were observed for the data which meet the following criteria: $M < 200$, $0.00 < P_d < 0.65$, and $0.16 < P_s < 0.20$. These data are represented in Table 5.2 as the shaded portion of the contingency table for deer. Expected values in these cells differed as much as $\pm 2000$ from observed cell frequencies. Positive deviations indicate observed frequncies are larger than expected

TABLE 5.2. Areas in the Contingency Table for Deer in Which Observed and Expected Cell Frequencies are the Most Similar.

| Probability of Success | Miles | Probability of Draw < 0.15 | 0.51–0.65 | 0.66–0.80 | 0.81–0.95 | >0 .95 |
|---|---|---|---|---|---|---|
| <0.10 | < 25 | | | | | |
| | 25– 50 | | | | | |
| | 51–100 | | | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | | |
| | 301–400 | | | | Small | |
| | > 400 | | | | Residuals | |
| 0.1–0.15 | < 25 | | | | | |
| | 25– 50 | | | | | |
| | 51–100 | | | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | | |
| | 301–400 | | | | Small | |
| | > 400 | | | | Residuals | |
| 0.16–0.15 | < 25 | | | | | |
| | 25– 50 | | Large | | | |
| | 51–100 | | Residuals | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | | |
| | 301–400 | | | | | |
| | > 400 | | | | | |
| 0.21–0.25 | < 25 | | | | | |
| | 25– 50 | | | | | |
| | 51–100 | | | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | | |
| | 301–400 | | | | Small | |
| | > 400 | | | | Residuals | |
| 0.26–0.30 | < 25 | | | | | |
| | 25– 50 | | | | | |
| | 51–100 | | | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | | |
| | 301–400 | | | | Small | |
| | > 400 | | | | Residuals | |
| > 0.30 | < 25 | | | | | |
| | 25– 50 | | | | | |
| | 51–100 | | | | | |
| | 101–200 | | | | | |
| | 201–300 | | | | Small | |
| | 301–400 | | | | Residuals | |
| | > 400 | | | | | |

frequencies. Likewise, negative residuals indicate observed frequencies are smaller than expected. The patterns associated with the signs of the residuals provided ambiguous information about the relationships among the data. Observed frequencies appear much larger and much smaller than expected cell values in a somewhat random fashion.

Residual analyses were also performed on the marginal subtables M versus $P_s$ & M versus $P_d$. Large residuals appeared to be closely linked with the miles variable over the range of 0–200 for all levels of $P_s$ and $P_d$. This result indicates that a breaking point of 200 miles traveled distinguishes two distinct "types" of hunt applicants.

The results of the residual analysis of the contingency table and marginal subtables indicate two things about hunt applicants' behavior. Hunters who are willing to travel less than 200 miles are a distinct group from those who are willing to travel more than 200 miles. Since the majority of observations fall in the M>200 categories, the applicants willing to hunt "close" to home represent an extreme when taken in the context of the study group as a whole. The applicants who fall in the most extreme residual categories would be those persons who wish to hunt "close" to home at a site which has a low probability of receiving a permit and only a mediocre probability of killing an animal. It seems reasonable to assume, and is indicated by the analysis that this type of individual may be considering other variables than just those in the contingency table in his decision to demand a hunting area. Further, where distance traveled is constrained to under 200 miles, the available choice sites decrease. This group of hunters has limited quality options and is not able to play the odds of receiving benefit as freely as others.

In Table 5.2 the hatched and cross hatched areas of the contingency table represent the data categories in which residuals are the lowest. Here the expected cell frequencies and the observed frequencies are most similar. Two distinct behavior types can be identified from the pattern of residuals. The hatched portion of the table represents applicants who desire a high probability of being drawn for a permit and a high probability of success of kill. The applicants are willing to travel more than 200 miles from their residence. These applicants represent the most serious of the possible hunter types and probably pay most attention to published probability odds. Because distance traveled is less of a constraining factor, these applicants are able to best play the odds of receiving benefits.

The cross-hatched portion of Table 5.2 represents those applicants who desire a high probability of draw and are willing to travel more than 200 miles for the hunting experience. To this group an actual kill plays a secondary role to the hunting experience as a recreational adventure. This type of applicant represents the second most serious type of hunter and perhaps the most serious type of recreator in the study group. It would be expected that this group of individuals along with the serious hunter group would fit the expected trends in the contingency tables better than most other groups.

A second method of examining trends in the data which contribute to a significant lack-of-fit of models is called partitioning for chi-squares. Marginal subtables of two variables each are created for each category of the third variable. The expected cell frequencies and Pearson's chi-square statistics are then calculated for each subtable. A small chi-square statistic for a two-variable subtable indicates that over the given range of the third excluded variable the data fits the table the best. Large chi-squares indicate the data contributes to a lack-of-fit of models.

The largest chi-square statistics occured over variable ranges that correspond to the largest residuals. Partitioning by chi-squares indicates that the data represented as the shaded area in Table 5.2 contributes to the significant lack-of-fit of models fitted to the contingency table. The results of the analysis of marginal subtables for the best fitting data were also similar to the results indicated by the residual analysis. And where success probabilities were low, observations in high probabilities of draw and high miles provided low chi-squares. Where draw probabilities were high, the observations in areas of the table with high probabilities of success and high miles contributed least to models' lack-of-fit. These observations fall in the hatched and cross hatched areas of Table 5.2. While Table 5.2 is used as an example for deer, patterns in the javelina data are quite similar. Conclusions about behavior patterns for deer hunters correlate almost exactly with those that can be drawn about javelina hunters by examining the results of the residual and marginal subtable analyses.

The marginal subtables which contributed least to a lack of models' fit were observed to have chi-square statistics of approximately one-third the magnitude of chi-squares associated with other marginal subtables. The smallest chi-square of 4,588 was associated with the subtable miles vs. draw at a level of success above 0.2. The chi-square statistics shows that even if this segment of data alone were tested, the model of independence would not provide an adequate fit for the data. Results of the subtable analysis confirm the results indicated in the initial stage of analysis when all models were tested. While variables are significant in explaining the data, specification of effects cannot be modeled simply.

## County-wide Models

In an attempt to test alternatives to the statewide model for hunting demand an analysis of hunter demand from individual county models was undertaken. Maricopa County, Pinal County, and Pima County were chosen first for the analyses because they include some of the most densely populated regions of the State (Arizona Statistical Review, 1979). It was thought that the quality of hunting sites, as represented by total numbers of site applicants, might be obscured by the uneven distribution of population throughout the state. Brown, Sorhus, Chou-Yang, and Richards (1983) pointed out that uneven population distribution leads to biased demand estimations when individual observations are used in the travel cost method.

As an example relevant to this study, suppose Site B is of greater quality than Site A. Site A lies 100 miles from a metropolitan area and receives 1,000 hunt applications. Site B lies 100 miles from a sparcely populated town and receives 10 applicants. When comparing the site demands of A and B over the 100 mile categories Site A would appear to be of greater quality because it receives 900 more applicants than does Site B. When the demand for sites is compared only among people living in a given metropolitan area, the ambiguity is eliminated.

The analysis of county models was undertaken using the same procedures as for the statewide model. Given the decision variables M, $P_s$ and $P_d$, no models provided a good fit. The best model was [$P_dM$, $P_dP_s$, $MP_s$] associated with javelina hunt applicants of Pima County. With 150 d.f. this model had a Peason chi-square statistic of 11,967. For this model to be nonsignificant and thus provide an adequate fit, the chi-square value needs to be less than 100. Results show that problems associated with disparities in population density throughout the state do not explain the lack-of-fit of the models to the data.

## Observed and Fitted Psuedo-Demand Functions

Due to the results of the models' screening, the objectives of determining fitted psuedo-demand functions for deer and javelina could not be met. Fitted psuedo-demand functions were to be read from the fitted frequencies of the contingency tables. Since the decision variables effects could not be modeled in any relevant manner, the fitted frequency table and fitted psuedo-demand functions could not be generated. Without the specification of appropriate site demand functions, the valuation of hunting resources is not possible. While the analysis was unable to define a simple mathematical function which relates the effects among the decision variables, examination of the observed contingency tables and observed psuedo-demand functions reveal important trends in these effects. Figures 5.1, 5.2, and 5.3 show the observed relationships for deer hunting on three separate areas.

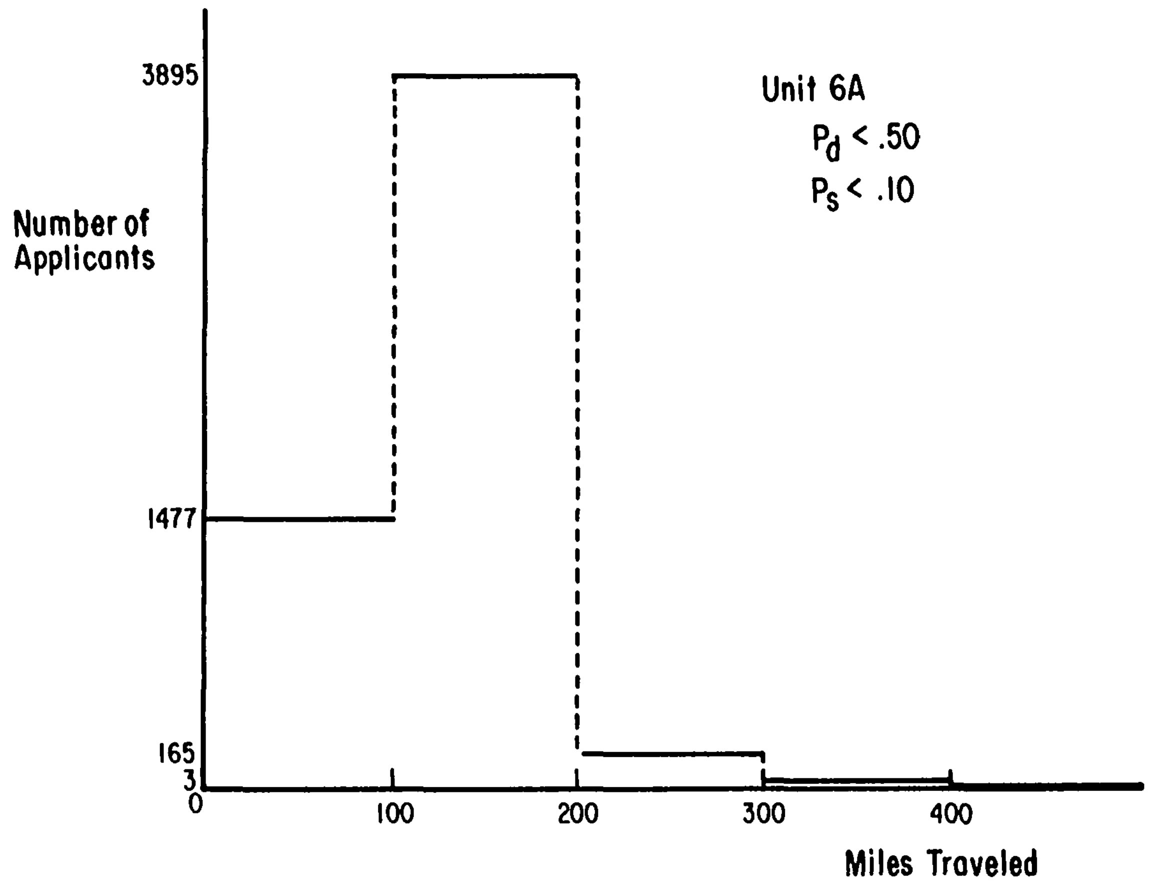


Figure 5.1. Relationship of number of applicants to miles traveled for a low quality deer hunting area.

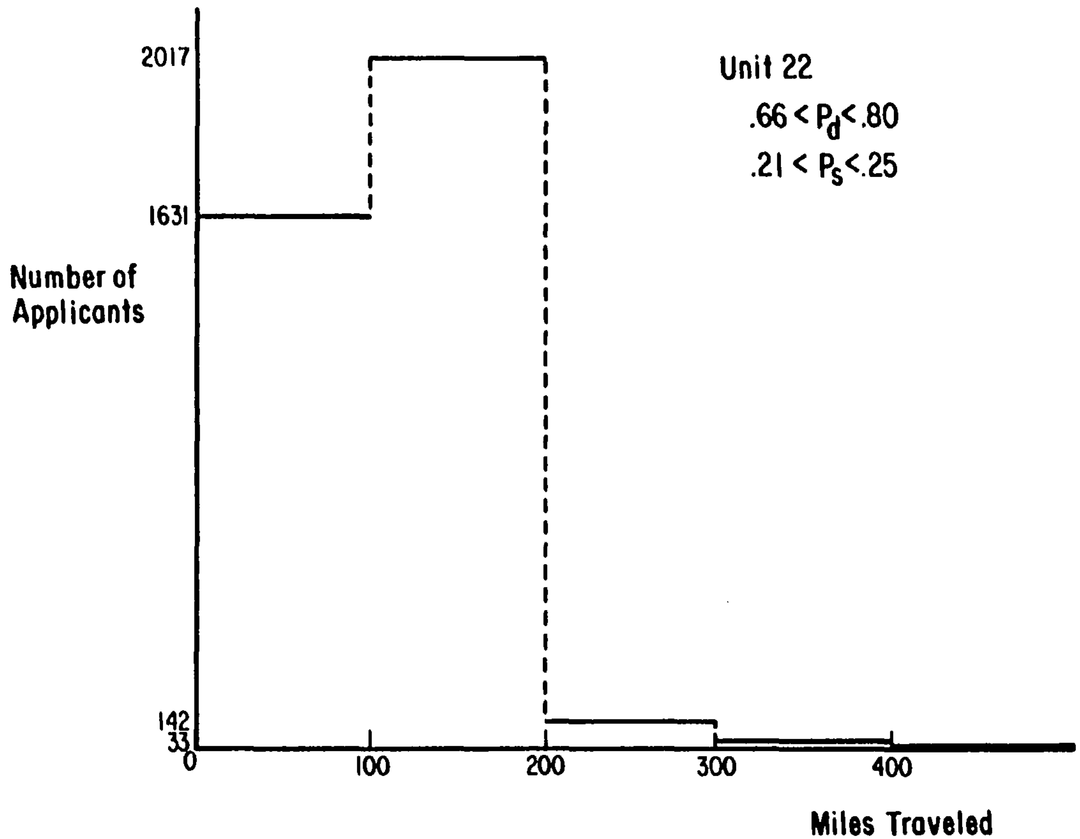


Figure 5.2. Relationship of number of applicants to miles traveled for a medium quality deer hunting area.

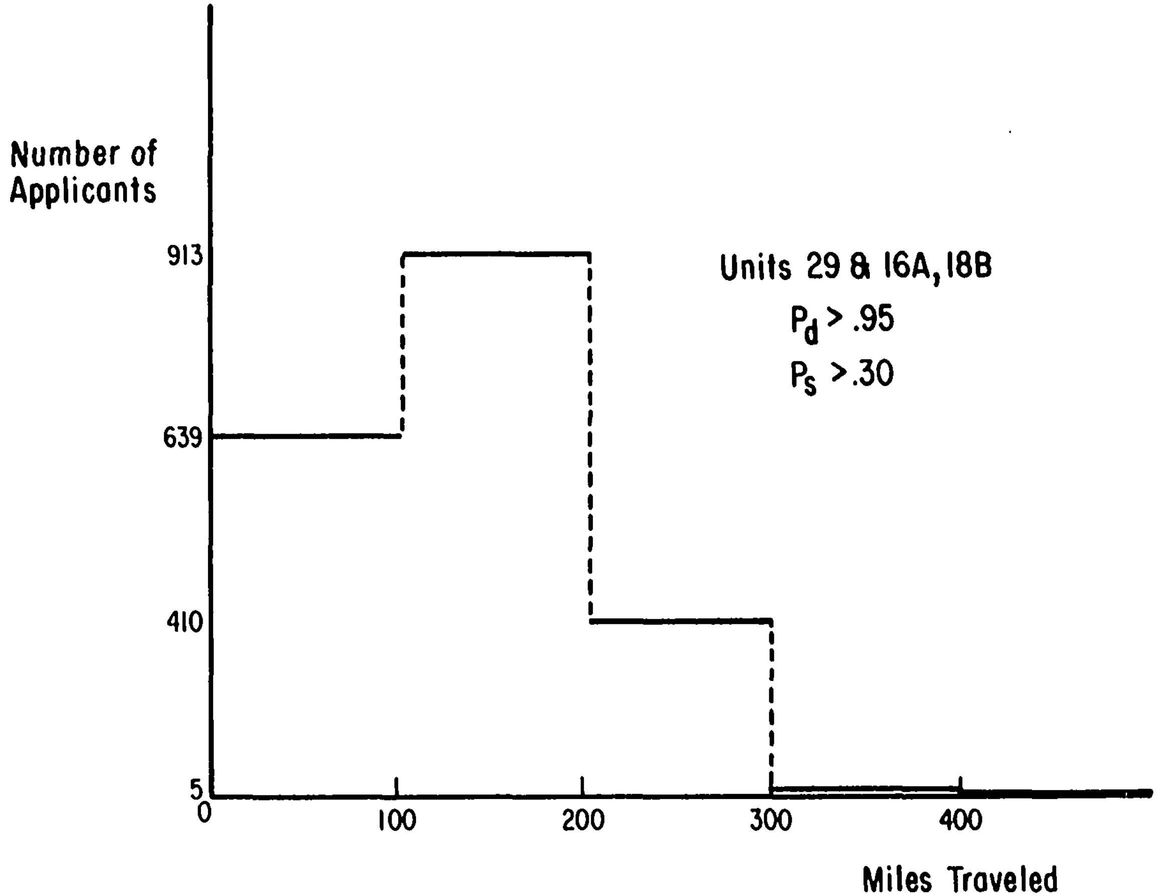


Figure 5.3. Relationship of number of applicants to miles traveled for high quality deer hunting areas.

Figures 5.1, 5.2, and 5.3 represent sites of low, medium, and high quality, respectively.

These sites were chosen because the characteristics of the relationships between participation and miles traveled best represent relationships found in the data as a whole for both deer and javelina. Applicant participation increases as the miles ranges increase from 0-100 to 101-200, reaches a maximum in the 101-200 mile range, and decreases over all other ranges of miles in all cases. If this relationship were represented by a continuous function, it would take the shape of a parabola skewed slightly to the left. Figure 5.4 depicts a continuous function as indicated by the observed data. Mathematically, the functional relationship over the relevant range would be of the cubic form $a + bx - cx^2 + dx^3$.

The effects that miles traveled and the combination of quality variables have on participation are seen to vary depending on the levels of the variables. In Figures 5.1 to 5.3, over the 0-100 miles range, increases in the quality variables are associated with increases in participation at first and then with decreasing participation. Over the 101-200 miles range increases in the quality combinations are associated with decreases in participation. Over the 0-100 mile range a majority of applicants demanded a medium quality site, while over the 101-200 mile range a majority of applicants demanded a low quality site. If the appropriate model for the data was the saturated model $[MP_sP_d]$, it would be expected that the interaction of probabilities and thus applicants' perceptions of quality would be dependent on the level of the miles traveled variable.

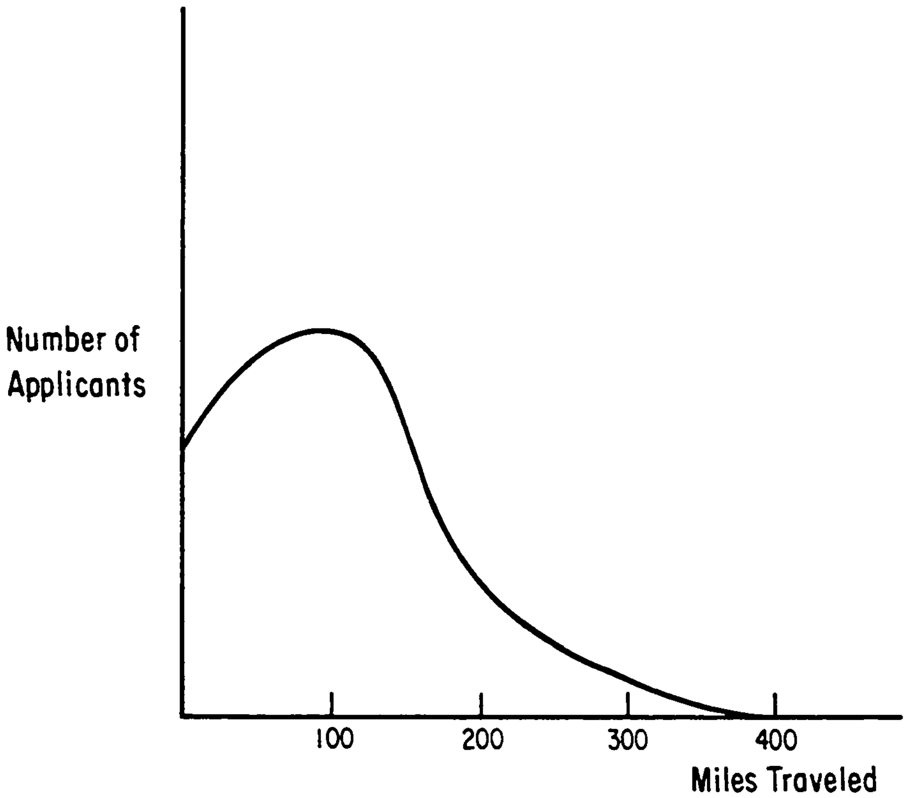


Figure 5.4. The hypothetical skewed parabola indicates a cubic function over the relevant range.

As previously discussed it is thought that part of the reason that participation among sites in the statewide model does not follow the hypothesis that for a given cost in miles applicants would demand the greatest quality of site is because totals of appliant particiation are observed by the inconsistent density of population throughout the state. To correct for this porblem one would need to express the number of applicants in terms of applicants per capita for each distance zone as suggested by Brown, et al. (1983). The definition of population density in distance zones centered about each applicant's residence would be at best complicated.

# CHAPTER 6

## CONCLUSIONS

In the analysis, two types of decision variables were hypothesized to effect hunt applicants' decisions to demand a particular hunting area. The first type of variable, miles traveled, was used as a proxy for the associated travel costs to the site. The second type of variable was associated with quality aspects of the area. Two quality variables, the probability of drawing an area permit and the probability of success of kill, were tested. The results of the analysis were that the hypotheses relating to these variables were incorrect. Travel cost and quality comparisons among areas play a significant role in determining hunting area demand. Results further indicate that the relationship among the variables is dynamically interactive and cannot be expressed in a simple log-linear model. When demand for and value of the resource cannot be mathematically quantified, conclusions about how policy decisions affect the various components in the demand for the resource and ultimately affect social well being cannot be drawn. Observations about how changes in the decision variables affect demand for a hunting site can only be made in a very general sense.

The justification for the inclusion of the miles traveled variable in this study was based on the travel cost theories as originally proposed by Clawson (1959, 1966) and Hotelling (1949), and extensively used for policy analysis since that time. Since the results of the log-linear analysis showed that miles traveled had significant effect in explaining site participation, comparison of the observed relationships with relationships used in the travel cost theories is warranted.

The deserved cubic relationship between participation and miles traveled, discussed in the last chapter, is consistent enough among sites to suggest implications contrary to those posited in the Clawson-Hotelling travel cost method. Central to the Clawson-Hotelling approach is the idea that the relationship between site participation and travel cost would be the same as hunters' reactions to varying site entrance fees. From demand theory we know that there exists an inverse relationship between price and quantity which gives demand curves their negative slope. Thus, we would expect that as entrance fees (or travel costs) rise, site participation would fall. In the Clawson-Hotelling approach a negatively sloped demand curve for the total recreation experience is estimated. Demand for the site is derived from the "total" curve by plotting miles traveled against per capita participation. The site demand curve, therefore, also retains its negative slope.

This analysis shows that over the range of 0-200 miles the relationship between desired participation and miles traveled is positive. Put in the framework of the Clawson-Hotelling approach, these results imply that as site entrance fees increase, so does the number of site participants who must travel less than 200 miles to the site. Clearly, when travel costs alone are considered in determining site demand, consumers appear to exhibit economically irrational behavior over near traveling distances.

Since it is more likely that it is the theories employed rather than consumers' behavior which is irrational, it is concluded that over near traveling distances, hunters consider other costs and benefits than those normally included in the travel cost method. For instance, part of the recreation experience includes traveling to less familiar environments and being out of touch with the everyday occurances of the home. These benefits tend to increase with the number of miles traveled away from a residence up to some point. Sublette and Martin (1975) gave empirical evidence that for many people benefits are derived from the activity of traveling in itself. Usually in the travel cost method, the act of traveling is considered neutral and benefits are derived at the point of arrival. These benefits play a part in determining site demand and must be included as part of the demand for the total experience. This analysis suggests, that at least under travel distances of up to 200 miles, that the Clawson-Hotelling two-stage method of site demand estimation does not accurately estimate Arizona hunters' actual behavior.

In 1980, Smith and Kopp suggested a test to define the outer bound or spatial limits of demand function parameters estimated by travel cost methods. Over large distance ranges, variances in costs and travel times associated with alternative forms of travel tend to obscure the significance of an estimated demand curve. From the results of this analysis, it is suggested that there also exists an inner spacial bound under which the demand estimation also loses meaning. Rather than the cost variance suggested by Smith and Kopp (1980), it is suggested here that variance in benefits associated with recreation activities near one's residence tends to decrease significance of demand parameters estimated by the traditional travel cost method. One should allow for positive as well as negative benefits from travel distance.

The difference betwenn applicants who are willing to travel less than 200 miles and those willing to travel more than 200 miles and those willing to travel more than 200 miles to a hunting area may be the difference between the way day hunters and overnight hunters perceive different benefits from an area. They also have different willingnesses to pay for the area because they are purchasing different commodities. Results of the analysis indicate that persons who are willing to travel less than 200 miles perceive different area benefits than those willing to travel more than 200 miles. Therefore, two separate demand curves for a given site must be examined. One demand for day hunters and one for overnight hunters.

The second set of conclusions from this analysis relate to aspects of site quality. Results indicated that the chosen quality variables were significant in determining applicant participation. Quality, at least in terms of the probability of success of being drawn and success of kill, appears to have bearing on where hunters want to hunt. In the introduction it was posited that applicants could use the published probabilities associated with site quality that are available through the Game and Fish Department. Results indicate that these probabilities are used as decision criteria for site demand in a dynamic game playing process which weighs perceived benefits against costs.

The fact that area quality is among the decision variables used in the choice of an area would have implications to area quality control policy. However, unless the relatonship among the quality variables and their effects on applicants can be defined and valued, the value of investments in area quality improvements cannot be estimated. Part of the dynamic game playing process of weighing the quality variables in the decision to choose a hunting area is dependent upon applicants' valuations of area quality in previous seasons. The probability of draw ($P_d$) is equal to the number of permits available divided by the number of applicants to the area. A vote in favor of an area's physical quality then would decrease the value of $P_d$ in the following year. High quality areas, in terms of the physical characteristics of the area, would be associated with low values of $P_d$. If, however, we define a high quality area in terms of the location which provides the greatest odds of receiving a permit and thus beneits of attendance, then high values of $P_d$ indicate high area quality. While the results of the analysis indicate that $P_d$ is highly significant in determing area demand, it is not clear how the probability of draw value is perceived and incorporated into the game playing process.

The general objective of this study was to test the applicability of log-linear modeling in the recreation framework. Results of the tests indicate mixed conclusions. The analysis appears most valuable in testing for significant variables in demand studies. Much insight into demand at specific sites can be gained from examining data in the contingency table framework. Further attributes of this type of modeling are difficult to ascertain because of the general lack of fit of the models tested. This result indicates that further research should be undertaken.

First, a method of incorporating per capita population data needs to be established when using individual observations in the anlaysis, as suggested by Brown et al., (1983). It may be necessary to examine only population centers in Arizona, rather than the statewide model to correct for ambiguities in the interpretation of the results. the total of applicants from Pima and Maricopa Counties accounts for 78 percent of the total number of applicants throughout the state for javelina hunts. Therefore, ambiguities due to population density problems occur in 22 percent of the data. It was seen from the countywide models that models' fit improved substantially over the statewide model. It is hoped that were individual observations put on a per capita basis, the fit of models would improve further. Since less than one-quarter of the data used in the statewide model contained possible population anomolies, conclusions about the relationship between miles traveled and number of applicants to a given quality of site can be made with a fair degree of certainty. Population adjustments incorporated into the analysis would be expected to clarify rather than redefine this relationship.

Secondly, further decision variables that affect hunting demand might be hypothesized and tested in the log-linear framework. An example would be to obtain income information about applicants and predict a variable relaying the time costs associated with travel. The interaction of this cost variable with the other posited decision variables may add information that leads models toward more independence and better fits.

It may be found that household decisions to demand a sight are affected by the availability of associated family activities, or that group decisions made by members of joint hunting applications affect the choice of site. These types of factors are usually considered as part of the demand for the total recreation experience. The log-linear method may be used to examine whether variables can be separated out of site demand as in the Clawson-Hotelling TCM or whether their presence is vital to explanations of site demand.

Lastly, it is suggested that data be collected on intended length of stay at the site for all applicants in the study group examined in this analysis. Applicants could then be subdivided into day hunters and overnight hunters, and separate demand curves for each group could be estimated.

## REFERENCES


Allen, P. G., T. H. Stevens, and S. A. Barret, "The Effects of Variable Omission in the Travel Cost Technique," Land Economics, 57 (May, 1981).

Arizona Department of Transportation, Arizona Highways 1981 Road Map. Phoenix, Arizona: Arizona Highways Magazine, 1981.

Arizona Game and Fish Commission, Arizona Game Survey and Harvest Data Summary. Phoenix, Arizona: Federal Aid Project W-53R-31, 1980.

Arizona Game and Fish Commission, Arizona Game Survey and Harvest Data Summary. Phoenix, Arizona: Federal Aid Project W-53R-31, 1981.

Arizona Game and Fish Commission, Arizona Game Survey and Harvest Data Summary. Phoenix, Arizona: Federal Aid Project W-53R-31, 1982.

Arizona Game and Fish Commission, "Secondary Data for Hunt Applicants in 1980-1981 Season," unpublished, 1982.

Bartlette, M. S. "Contingency Table Interactions," Journal of Royal Statistics Society Supplemental, 2(1935):248-252.

Beazley, R. I. "Some Considerations for Optimizing Public Forest Recreational Development and Value," Journal of Forestry, 59(1961):644-650.

Benedetti, J. K., and M. B. Brown, "Alternate Method of Building Log-Linear Models," Proceedings of the 9th International Biometric Conference, 2(1976):209-227.

Bishop, Y. M. M., S. E. Feinberg, and P. W. Holland, Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press, 1975.

Blalock, H. M., Jr., Social Statistics, 2d ed. New York: McGraw-Hill, 1972.

Brown, W. G., Ajmer Singh, and E. N. Castle, "An Evaluation of the Oregon Salmon and Steelhead Sport Fishery," Technical Bulletin No. 78, Agricultural Experiment Station, Oregon State University, Corvallis, 1964.

Brown, W. G., and Farid Nawas, "Effects of Aggregation Upon the Evaluation and Specficiation of Outdoor Recreation Demand Functions," Proceedings of the Forty-Fifth Annual Meetings of the Western Agricultural Economics Association, Logan, Utah, 1972.

Brown, W. G., Colin Sorhus, Bih-lian Chou-Yang, and J. A. Richards. "Using Individual Observations to Estimate Recreation Demand Functions: A Caution," American Journal of Agricultural Economics, 65(February, 1983):154-159.

Capel, C. F., and R. K. Pandry, "Estimating the Benefits from Deer and Moose Hunting in Manitoba," Canadian Journal of Agricultural Economics, 2(July,1972):7-16.

Clawson, Marion, Methods of Measuring the Demand for and Value of Outdoor Recreation, Resources for the Future, Reprint No. 10, Washington, D.C., 1959.

Clawson, Marion, and J. L. Knetsch, Economics of Outdoor Recreation. Baltimore: John Hopkins Press, 1966.

Cox, D. R. "The Continuity Correction." Biometrika, 57(1970):217-219.

Currie, J. M., J. A. Murphy, and A. Schmitz, "The Concept of Economic Surplus and Its Use in Analysis," The Economic Journal, 81(May,1971):741-799.

Dempster, A. P., "An Overview of Multivariate Analysis," Journal of Multivariate Analysis, 1(1971):316-346.

Dixon, W. J., and M. B. Brown, eds, Biomedical Computer Programs P-Series 1981, 2d ed. Berkeley, California: University of California Press, 1981.

Dupuit, J., "On the Measurement of the Utility of Public Works." Annals des Ponts et Chaussees, second series, vol. 8, 1844; Translation Reprinted in D'Munby, Transport, 1968.

Feinberg, S. E., "The Analysis of Incomplete Multiway Contingency Tables," Biometrics, 28(1972): 177-202.

Feinberg, S. E., The Analysis of Cross-Classified Categorical Data, 2d ed., Cambridge, Mass: MIT Press, 1981.

Fisher, R. A., "The Use of Multiple Measures in Taxonomic Problems," Ann. Eugen., 7(1936):179-188.

Gum, R. L., and W. E. Martin, "Problems and Solutions in Estimating the Demand for and Value of Rural Outdoor Recreation," American Journal of Agricultural Economics, 57(November, 1975):558-566.

Helfrich, M. A., "Towards Estimating the Optimum Combination of Cattle and Elk on the Apache-Sitgraves National Forest of Arizona," Masters thesis, University of Arizona, 1981.

Hicks, J. R., "The Four Consumers' Surpluses," Review of Economic Studies, 11(1943):31-41.

Hotelling, Harold, "The Economics of Public Recreation," The Prewitt Report, mimeographed, Washington: Land and Recreation Planning Divison, National Park Service, 1949.

Killian, R. A., and D. A. Zahn, "A Bibliography of Contingency Table Literature: 1900-1974," International Statistics Review, 4(1976):71-112.

Marshall, A., Principles of Economics, eighth ed. London: Macmillan and Co., limited, 1936.

Martin, W. E., R. L. Gum, and A. H. Smith, "Values for Hunting, Fishing, and General Rural Outdoor Recreation in Arizona for 1970." Technical Bulletin No. 211, Agricultural Experiment Station, University of Arizona, Tucson, June, 1974.

Martin, W. E., J. C. Tinney, and R. L. Gum, "A Welfare Economic Analysis of the Potential Competition Between Hunting and Cattle Ranching," Western Journal of Agricultural Economics, 3(1978):87-98.

McConnell, K. E., and V. A. Duff, "Estimating Net Benefits of Outdoor Recreation Under Conditions of Excess Demand," Journal of Environmental Economics and Management, 2(February, 1976):224-230.

Outdoor Recreation Resources Review Commission. Public Outdoor Recreation Areas-Acreage, Use Potential, ORRRC Study Report No. 1, Washington, D. C., 1962.

Outdoor Recreation Resources Review Commission. Participation in Outdoor Recreation: Factors Affecting Demand Among American Adults, ORRRC Study Report No. 20, Washington, D. C., 1962.

Outdoor Recreation Resources Review Commission, Trends in American Living and Outdoor Recreation, ORRRC Study Report No. 22, Washington, D.C., 1962.

Outdoor Recreation Review Commission, Economic Studies in Outdoor Recreation, ORRRC Study Report No. 24, Washington, D.C., 1962.

Outdoor Recreation Review Commission, Prospective Demand for Outdoor Recreation, ORRRC Study Report No. 26, Washington, D.C., 1962.

Plackett, R., The Analysis of Categorical Data, New York: Hafner Press, 1974.

Prewitt, R. A. The Economics of Public Recreation: An Economic Study of the Monetary Evaluation of Recreation in the National Parks, U.S. Department of Interior, National Parks Service, Washington, D.C., 1949.

Randall, Alan, "Evaluating Non-Market Goods and Services: Some Conceptual Considerations," Staff Paper No. 51, University of Kentucky, Department of Agricultural Economics, July, 1977.

Rao, C. R., Linear Statistical Inference and Its Applications, 2d ed., New York: John Wiley and Sons, 1973.

Smith, V. K., and R. J. Kopp, "Spatial Limits of the Travel Cost Recreation Demand Model," Land Economics, 56(February, 1980):64-72.

Sublette, W. J., and W. E. Martin, "Outdoor Recreation in the Salt Verde Basin of Central Arizona: Demand and Value," Technical Bulletin No. 218, Agricultural Experiment Station, University of Arizona, Tucson, June 1975.

United States Postal Service, Zip Code of Cities and Communities in Arizona Sherman Oaks, California: Western Research Co., 1979.

U. S. Water Resources Council, "Procedures for Evaluation of National Economic Development," Federal Register, 24 May, 1979.

U. S. Water Resources Council, "Principles and Standards for Planning Water and Related Land Resources," Federal Register, 14 December, 1979: 72892-72975.

U. S. Water Resources Council, "Proposed Rules: Principles, Standards, and Procedures for Planning Water and Related Land Resources," Federal Register, 14 April, 1980:25302-25348.

Valley National Bank of Arizona, Arizona Statistical Review, 35th annual edition, Phoenix, Arizona: Economic Research Department, September, 1979.

Ward, Frank, "A Review of Problems and Solutions in the Use of Travel Cost Method for Valuing Recreation Resources," Staff Report No. 14, New Mexico State University, Department of Agricultural Economics and Agricultural Business, January, 1980.

Wennergren, Boyd, "Surrogate Pricing of Outdoor Recreation," Land Economics, 43(February, 1967a):112-116.

Wennergren, Boyd, "Demand Estimates and Resource Values for Resident Deer Hunting in Utah," Bulletin No. 469, Agricultural Experiment Station, Utah State University, Logan, 1967b.

Wetzell, J. N., "Estimating the Benefits of Recreation Under Conditions of Congestion," Journal of Environmental Economics and Management, 4(September, 1977):239-246.

Willig, R. D., "Consumer's Surplus Without Apology," The American Economic Review, 66(1976): 589-597.

Yates, F., "Contingency Tables Involving Small Numbers and the $X^2$-Test," Journal of Royal Statistics Society Supplimental, 1(1934):217-235.