NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

PREDICTING COAST GUARD ENLISTED ATTRITION

by

Douglas Allen Blakemore

March 1992

Thesis Advisor: Robert R. Read

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY 2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution is unlimited.
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Naval Postgraduate School 6b. OFFICE SYMBOL OR	7a. NAME OF MONITORING ORGANIZATION
6c. ADDRESS (City, State, and ZIP Code) Monterey, CA 93943-5000	7b. ADDRESS (City, State, and ZIP Code)
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO.

11. TITLE (Including Security Classification)
PREDICTING COAST GUARD ENLISTED ATTRITION

12 PERSONAL AUTHOR(S)
BLAKEMORE, Douglas Allen

13 TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. Page Count
Master's thesis	FROM TO	1992, March	59

16. SUPPLEMENTAL NOTATION
The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17. COSATI CODES

FIELD	GROUP	SUB-GROUP

18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
Attrition, survival analysis, Weibull regression model, Binomial counting model, survivor function, accelerated failure time model, Coast Guard Attrition

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

This study examines enlisted attrition behavior for the U.S. Coast Guard and develops a model that projects attrition figures. Survival analysis techniques are used to analyze the empirical attrition behaviors associated with an individuals sex, race, marital status, and military occupational skill (MOS). In this study males tend to have higher survival probabilities than females, non-caucasians higher than caucasians, and married persons higher than those not married. Aviation MOSs have the highest survival probabilities and technical MOSs have the lowest.

Modelled survivor functions are developed for two paygrades because each contain small personnel inventories. These modelled survivor functions do not fit the data as well as desired but are nonetheless used pending the development of sharper alternatives.

Finally, a counting model based on the Binomial Distribution is developed that projects monthly enlisted attrition figures.

20 DISTRIBUTION/AVAILABILTIY OF ABSTRACT [X] UNCLASSIFIED/UNLIMITED [ ] SAME AS RPT. [ ] DTIC	1a. REPORT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL Robert R. Read	22b. TELEPHONE (Include Area Code) (408)646-2382 22c. OFFICE SYMBOL OR/Re

DD Form 1473, JUN 86 Previous editions are obsolete. SECURITY CLASSIFICATION OF THIS PAGE
S/N 0102-LF-014-6603 Unclassified

Predicting Coast Guard Enlisted Attrition

by

Douglas Allen Blakemore
Lieutenant, United States Coast Guard
B.S., U.S. Coast Guard Academy

Submitted in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS ANALYSIS

from the

NAVAL POSTGRADUATE SCHOOL
March 1992

Author: Douglas Allen Blakemore

Approved by: R.R. Read, Thesis Advisor

Lyn R. Whitaker, Second Reader

Peter Purdue, Chairman
Department of Operations Research

ABSTRACT

This study examines enlisted attrition behavior for the U.S. Coast Guard and develops a model that projects attrition figures. Survival analysis techniques are used to analyze the empirical attrition behaviors associated with an individuals sex, race, marital status, and military occupational skill (MOS). In this study males tend to have higher survival probabilities than females, non-caucasians higher than caucasians, and married persons higher than those not married. Aviation MOSs have the highest survival probabilities and technical MOSs have the lowest.

Modelled survivor functions are developed for two paygrades because each contain small personnel inventories. These modelled survivor functions do not fit the data as well as desired but are nonetheless used pending the development of sharper alternatives.

Finally a counting model based on the Binomial Distribution is developed that projects monthly enlisted attrition figures.

TABLE OF CONTENTS

I. INTRODUCTION

A. PROBLEM STATEMENT

The United States Coast Guard (CG) is a military service operating within the Department of Transportation. Its main missions are aids to navigation, boating safety, defense operations, maritime law enforcement, environmental response, marine inspection, ice operations, marine licensing, marine science, search and rescue, port security and safety, and waterways management. The CG maintains an active duty hierarchical personnel structure with approximately 6,000 officers and 30,000 enlisted members. The CG Personnel Workforce Planning (PWP) office located in Washington, D.C. has the overall responsibility for projecting the active duty enlisted personnel needs of the CG. Their responsibilities include the calculating of personnel stocks, promotion requirements, attritions, and recruitment needs. The current method for predicting monthly enlisted attrition by paygrade is to calculate a running average or mean of past monthly attrition and use this mean as a forecast of future attrition. This forecasting method is crude, ignores many issues, and is not accurate enough to meet the needs of the CG.

Why does a person separate from the CG? What are the characteristics associated with this person? Are these characteristics the same for each paygrade? Previous CG studies have focussed on the first question but have not looked at attrition data analytically. The first question will not be addressed. This thesis will look at the next two questions and will propose a method, based on a survival analysis of CG enlisted attrition data, for calculating monthly personnel attrition. This study is a first attempt at examining CG enlisted attrition using data analysis tools. It is designed to provide a procedure for studying attrition, and for developing a basic model that calculates monthly attrition which can be expanded by future studies.

B. COAST GUARD ENLISTED PERSONNEL BACKGROUND INFORMATION

The Coast Guard enlisted personnel structure is composed of 9 paygrades, E-1 to E-9, lowest to highest respectively. (There is an E-10 position filled by one person but this thesis does not address this paygrade.) Paygrade E-1 is composed of persons attending CG basic training school and paygrade E-2 is composed of persons who have completed basic training school, have been assigned to active duty commands, and are preparing to attend a CG specialty school (also called A-school). Paygrades E-3 through E-9 are divided into 24 active duty military occupational skills (MOS) or sub-specialties listed below:

- Aviation Machinist's Mate (AD)
- Aviation Electrician's Mate (AE)
- Aviation Structural Mechanic (AM)
- Aviation Survivalman (ASM)
- Aviation Electronic Technician (AT)
- Boatswain's Mate (BM)
- Damage Controlman (DC)
- Electrician's Mate (EM)
- Electronics Technician (ET)
- Fire Control Technician (FT)
- Gunner's Mate (GM)
- Health Services Technician (HS)
- Machinery Technician (MK)
- Musician (MU)
- Marine Science Technician (MST)
- Public Affairs Specialist (PA)
- Quartermaster (QM)
- Radarman (RD)
- Radioman (RM)
- Storekeeper (SK)
- Subsistence Specialist (SS)
- Sonar Technician (ST)
- Telephone Technician (TT)
- Yeoman (YN)

An enlisted man (EM) receives an MOS after either completing a specialty school or by completing an intensive on-the-job training program. Both methods are usually completed while the EM is in the E-3 paygrade. An E-3 who has not obtained an MOS is called nonrated; E-3's through E-9's hold only one MOS at a time and usually keep that MOS throughout their careers.

Promotions to the next highest paygrade are performed by MOS and are dictated by the needs of the CG. Promotions to paygrades E-5 through E-9 are vacancy driven and promotions to paygrades E-3 and E-4 occur on qualification for advancement. Promotion to paygrade E-2 occurs upon completion of basic training.

Enlisted personnel sign service obligation contracts (enlistments) that require the person to serve in the CG for a set number of years. Enlistments usually last four years. (There have been two year enlistments but these have been terminated.) Upon completion of an enlistment and upon approval by the CG, a person may sign a new contract (reenlist), or separate from the CG. Enlisted attrition generally occurs because of:

- retirement - minimum 20 years active duty service,
- non-reenlistments - an EM chooses not to reenlist or the CG chooses not to reenlist the person,

- administrative reasons - an EM may depart the CG prior to the end of his/her enlistment "for the convenience of the government",

- death or disability,

- selection to an officer program.

For the purposes of this study the term separation refers to a person leaving the CG for any of the above reasons.

C. PROCEDURE

This thesis examines, by paygrade, individual personnel characteristics associated with attrition and develops a model that calculates monthly attrition. Survival analysis techniques are used to study behaviors associated with these attrition characteristics. Survival analysis is a data analysis technique that is useful in examining failure time data. Many of the visible applications have been in medical research where a typical "failure" is the death of a patient who had been subjected to some type of treatment. Obviously it would not be practical to wait for all the patients to "fail" before analyzing the data. It is more convenient to end the experiment at some specified time and do an analysis of all the data. The observables that did not "fail" by the end of the data collection period are called censored observations. They must be treated differently since their failure times are yet to occur and are hence unknown. Survival analysis allows for censored failure times and provides techniques for including them in the analysis. The primary tool used in the survival analysis will b the estimation of survivor functions or curves. These curves provide probabilities that a person survives up to and beyond a set time, t. Their estimators are often called product limit estimators because of the technique used. Survival probabilities are the compliment of attrition rates and are needed to calculate expected monthly attrition figures.

The personnel data are cross classified by paygrade and each paygrade is partitioned by certain individual personnel characteristics. Paygrades E-1 and E-2 will be partitioned by the characteristics of sex, marital status, and race. Paygrade E-3 will be partitioned by these same characteristics and by whether the person had obtained an MOS or was nonrated. Paygrades E-4 through E-9 will be partitioned by the same characteristics as E-1s and E-2s and by CG MOS. MOSs will be grouped according to categories listed in Table 1 on page 7.

TABLE 1. MOS GROUPINGS

Group	Rates
Aviation	AD, AE, AM, ASM, AT
Administrative	YN, SK
Technical	EM, ET, FT, MST, ST, TT
Skilled	HS, SS, MU, PA
CG Professional	BM, DC, GM, MK, QM, RD, RM

MOSs are grouped for two reasons. MOSs within the same group have similarities that should cause them to exhibit common attrition behavior and hence have common survival probability values. The MOSs within each category all have several common attributes; specifically, the type of training needed to acquire the MOS, the type of duty or jobs associated with those MOSs, and the possible civilian employment opportunities available for persons with these MOSs. MOSs are also grouped to increase the personnel inventory within each paygrade required for a survival analysis. Chapter II contains an explanation on these MOS groups and examines the data used in this study.

For most MOSs stocks are large and empirical survival probabilities can be used. Modelled survival probabilities are needed if personnel stocks are smaller than preferred. Small personnel stocks are subject to large data variability; models are developed to provide stability. Paygrades E-8 and

E-9 could be considered too small to use empirical survival probabilities, thus regression models are examined in Chapter IV for these paygrades.

A model that projects an expected number (and variance) of monthly attritions will be developed and demonstrated in Chapter V. This model is based on principles of the Binomial distribution. Paygrades will be partitioned several times to produce multiple binomial experiments from which expected values and variances can be calculated. The product limit estimates generated by the survival analysis will correspond to the probabilities of success associated with a binomial experiment.

Finally conclusions and recommendations are contained in Chapter VI. Since this is the first time in recent years that enlisted attrition has been viewed using operations analysis tools, most of the conclusions should be unique. It is expected that additional studies utilizing more refined data will be needed to better understand, analyze, and predict CG enlisted attrition.

II. DATA OVERVIEW

A. POPULATION

Data for this thesis were provided by the CG Pay and Personnel Center located in Topeka, Kansas. The data were provided in two sets. The first set contained individual personnel records for all persons who separated during FY 91 (1 November 1990 - 31 October 1991). The second set contained individual personnel records for all enlisted persons who were on active duty the last day of FY 91 (31 October 1991). This second data set is in a sense a "snapshot" of the GC EM composition on 31 October 1991. In both data sets each personnel record contained the persons *sex*, *MOS* (if applicable), *rank*, *marital status*, *ethnic origin*, *date of current enlistment*, and *date of separation* (if applicable). Each persons length of service (in months) can be calculated from the data entries *date of current enlistment* and *date of separation*. It should be noted that the entry, *date of current enlistment*, did not necessarily correspond to when the person first joined the CG. This affects the calculation of a persons time in service. In certain cases the CG allows an EM to separate from the CG and then allows the person to rejoin the CG in the future at his or her past paygrade. Such persons' past service time is not ignored. The person maintains all accumulated service time for pay and retirement purposes. This condition was not reflected in the data provided for this study. There are numerous records of senior EMs having relatively little time in service when in all likelihood they have quite a bit of service time. This situation is most prominent in the survival analysis results of Chapter III. Appendix A contains procedures and Fortran programs needed to access and use the data.

B. LIMITATIONS

Analyzing only one years amount of personnel data presents two problems. The first and foremost is that the data lack temporal stability. More than one years worth of data is required to provide stability. Attrition study conclusions and modelling results based on unstable data are not as reliable as if they were based on time tested data. Secondly, the data may not have been as reliable as preferred because data inputs could not be verified. The FY 91 data were placed in storage by the Coast Guard PPC and were not accessible during the study. Recommendations contained in Chapter VI address these shortcomings further.

C. MOS GROUPS

As stated in Chapter I MOS groups were formed because of common attributes associated within each group. These attributes are listed in this section.

1. Aviation

Individuals with aviation MOSs have common A-school characteristics and common duty stations. (The CG has three basic types of duty stations: ashore, afloat, and air stations.) The aviation A-schools (specialty school) waiting lists are relatively long compared to the other MOS groups. This lengthens a persons service time in paygrade E-3, thus increasing the amount of time needed for promotion to E-4. Unlike many MOSs, individuals within this group are limited in their types of duty stations. They spend most of their entire career at CG air stations.

2. Administrative

These MOSs are again grouped because of their A-school characteristics, choice of duty locations, and possible civilian employment opportunities. The administrative A-schools have short waiting lists and short school durations. Promotion to E-4 in this MOS group is relatively fast compared to other groups. Administrative persons also can serve in almost any CG duty station or command. Since these persons skills are mainly clerical in nature it is felt that their civilian employment opportunities are limited.

3. Technical

Persons with technical MOSs share the same type of characteristics as the administrative MOS group, but with two opposite attributes. Their A-schools have long waiting lists, are long in duration, and have low graduation rates (as compared to the other MOS group A-schools). Like the administrative group, persons in this group can serve in almost any CG command. However, unlike the above group, technically grouped persons probably have more civilian job opportunities.

4. Skilled

This group is characterized by their scope of possible civilian employment opportunities. Persons within this group probably have marketable skills, possibly not as marketable as the technical group but more marketable than any other rate group.

5. CG Professional

This group is characterized by MOS obtainment and employment opportunities. As discussed in Chapter I, MOSs are obtained by graduating from an A-school or by completing an on-the job training program. This group is the only MOS group that incorporates on-the-job training programs for a number of their MOSs. MOSs are also designed for CG specific functions (ie. shipboard machinists, CG weapons technicians, Shipboard navigators, etc.). These skills have limited value for civilian employment.

D. SUMMARY OF PERSONNEL STOCKS

Table 2 provides a summary of personnel stocks partitioned by the rank, sex, race, and marital status for all CG EMs who were on active duty as of 31 October 1991.

TABLE 2. ACTIVE DUTY PERSONNEL STOCKS FY 91

Pay-grade	Total	Male	Female	Cauc.	Non-Cauc.	Mar.	Not Mar.
E-1	732	622	110	595	137	68	664
E-2	3383	3027	356	2776	607	622	2761
E-3	3651	3358	293	3031	620	1256	2395
E-4	7064	6364	700	6043	1021	3641	3423
E-5	5904	5410	494	4830	1074	4278	1626
E-6	5748	5427	321	4878	870	4805	943
E-7	2764	2696	68	2436	328	2436	328
E-8	569	564	5	501	68	517	52
E-9	289	289	0	256	33	270	19
Total	30,104	27,757	2347	25,346	4758	17,893	12,211

Note that race is divided into two categories: caucasian and non-caucasian. There are relatively few (as compared to the total EM population) minorities in the CG. All minorities were grouped to increase the personnel stocks within the race category. As can be seen the CG is a male, caucasian dominated service. Table 3 on page 14 provides the same statistics for Coast Guardsmen who separated during FY 91.

TABLE 3. PERSONNEL STOCKS FOR PERSONS SEPARATING DURING FY 91

Pay-grade	Total	Male	Female	Cauc.	Non Cauc.	Married	Not Married
E-1	779	637	142	601	178	83	696
E-2	522	438	84	429	93	115	407
E-3	831	731	100	712	119	276	555
E-4	1248	1094	154	1079	169	602	646
E-5	438	398	40	402	36	281	157
E-6	385	369	16	348	37	317	68
E-7	315	305	10	271	44	277	38
E-8	79	77	2	73	6	73	6
E-9	46	46	0	45	1	39	7
Total	4643	4095	548	3960	683	2063	2850

Approximately the same proportions of males to females, caucasians to non-caucasians, and married persons to persons not married separated from the CG as did those who did not separate during FY 91.

Table 4 on page 15 provides on a summary of personnel stocks for FY 91 active duty personnel by MOS group.

TABLE 4. FY 91 ACTIVE DUTY PERSONNEL STOCKS BY MOS GROUP

Paygrade	Aviation	Admin.	Technical	Skilled	CG Pro.
E-3	20	41	84	71	333
E-4	1142	970	1031	715	3206
E-5	921	949	781	490	2763
E-6	748	899	837	527	2737
E-7	262	353	452	260	1437
E-8	77	88	98	55	251
E-9	43	40	38	18	150
Total	3213	3340	3321	2136	10,877

The numbers in paygrade E-3 are small because not all E-3s have an MOS. Overall approximately 48% of all personnel belong to the CG professional MOS group. About 14% belong in the aviation group; 15% belong in the administrative group; 14% belong in the technical group; and 9% belong in the skilled rate group. Percentages within each paygrade are roughly the same. Table 5 on page 16 provides the same statistics for personnel who separated during FY 91.

TABLE 5. FY 91 SEPARATION PERSONNEL STOCKS BY MOS GROUP

Paygrade	Aviation	Admin.	Technical	Skilled	CG Pro.
E-3	1	5	12	23	65
E-4	75	157	180	100	733
E-5	26	49	91	26	246
E-6	31	46	51	36	221
E-7	22	51	23	26	193
E-8	11	15	14	8	31
E-9	3	6	3	3	31
Total	169	328	373	221	1518

E. DATA PARTITIONING

The data was partitioned by rank to form 9 paygrades. Each paygrade was further divided by the personnel characteristics sex, race, marital status, and MOS groups listed in Table 1 on page 7.

1. Paygrades E-1 and E-2

These paygrades were divided by the sex, race, and the marital status of the individual. Marital status indicated whether or not a person was currently married.

2. Paygrade E-3

In addition to the above, this paygrade was partitioned by the obtainment of an MOS. The person either had obtained one of the MOSs listed in Chapter I or had not (nonrated).

3. Paygrades E-4 through E-6

These paygrades were also partitioned by sex, race, and marital status. Additionally they were partitioned by MOS groups.

4. Paygrades E-7 through E-9

These paygrades were divided by race, marital status, and MOS group. Sex was not examined because of the scarcity of females in these paygrades.

III. SURVIVAL ANALYSIS

A. INTRODUCTION

As stated in Chapter I, survival analysis provides a method for analyzing censored failure time data. Failure times in this study refer to the number of months a person serves on active duty prior to separation. Failure times are assumed to be discrete and are rounded up to the next whole month figure. For example, if a person separates after 31 months, six days of service his or her failure time is 32 months. This chapter reviews the theory of survival analysis and graphically compares survivor functions of individual paygrades that have been partitioned by single personnel characteristics. Appendix C contains the SAS LIFETEST procedure and Fortran code required to estimate the survivor functions.

B. COMPUTATION OF SURVIVOR FUNCTIONS

Let T be a nonnegative random variable representing the separation time of an EM from a homogeneous population. The survivor function is defined to be the probability that T is at least as great as some time, t,

$$S(t)=Pr(T\geq t);\ 0\leq t<\infty. \qquad (3.1)$$

From a simple random sample, this can be estimated by

$$\hat{S}(t)=1-\left(\frac{\textit{number of observations with failure time} \leq t}{\textit{total number of observations}}\right). \qquad (3.2)$$

If T is a discrete random variable taking values $x_1 < x_2 \ldots < x_j$, with associated probability function $f(x_i) = Pr(T=x_i)$, then

$$S(t)=\sum_{i=1}^{j} f(x_i)H(x_i-t), \qquad (3.3)$$

where H is the Heaviside function, $H(x)=\{0 \text{ if } x<0;\ 1 \text{ if } x\geq 0\}$ [Ref. 1:pp. 5-7].

The presence of censored data, however, weakens the estimation of the survivor function in the right tail. All that can be safely assumed about an observation censored at time t is that the unobserved failure time is greater than t. This condition is relieved by using an alternate formulation of the survivor function using hazard rates. For discrete random variables, the hazard rate, λ_t, is the probability that the observation will immediately fail at time t given is has survived to time t. In the discrete case this can be estimated by

$$\lambda_t = \frac{\text{number of observations that fail at } t}{\text{number of observations with failure time} > t} . \quad (3.4)$$

Using hazard rates [Ref. 3], the survivor function may be expressed as

$$S(t) = \prod_{i: x_i \le t} (1-\lambda_i) . \qquad (3.5)$$

where λ_i is the hazard rate at x_i for $i=1,\ldots,j$. Estimating λ_j would allow an estimation of the survivor function, S(t). Chapter 2 of Reference 1 contains a detailed explanation of the computation of the estimated survivor function.

C. SURVIVAL ANALYSIS RESULTS

In all paygrades the marital status characteristic and sex characteristic behaviors were identical; married persons appeared to have a higher probability of survival than those who were not married, and males tended to have higher survival probabilities than did females. (See Figures 1 and 2 on page 40.) Non-caucasians in paygrades E-3 through E-9 had higher survival probabilities than did caucasians. Results for other characteristics are given below.

1. E-1

Caucasians and non-caucasians have about the same survivor functions until about two months. After two months caucasians tend to have higher survival probabilities. (See Figure 3 on page 41.)

2. E-2

Caucasians had higher survival probabilities than did non-caucasians up to about 20 months. (See Figure 4 on page 41.) After that non-caucasians had higher survivor probabilities. Also note the significant drop in probabilities at 24 months. The CG for a brief time allowed two year enlistments. This drop corresponds to enlisted persons separating at the end of a 2 year enlistment.

3. E-3

Individuals possessing an MOS had higher survival probabilities than did nonrates after 22 months. (See Figure 5 on page 42.)

4. E-4 and E-5

The aviation MOS group in both paygrades had the highest survival probabilities and the technical group had the lowest. The remaining three groups survival functions fell in between the aviation and technical survivor functions. (See Figure 6 on page 42.)

5. E-6

All MOS groups in this paygrade were very similar as shown in Figure 6 on page 42. All had very high probabilities of survival and all were very similar in shape. Again the aviation group had the highest survival probabilities.

6. E-7, E-8, and E-9

The survivor functions in all three paygrades appeared fairly flat with slightly decreasing slopes until about 220 months. (See Figure 7 on page 43.) After 220 months all survivor functions decreased at about the same rate. All survivor functions within these paygrades are characterized by high survival probabilities with the aviation group having the highest probabilities of survival.

IV. MODELLED SURVIVOR FUNCTIONS

A. INTRODUCTION

Regression models are typically used to model survival experiences from homogeneous populations [Ref. 1]. These models allow for failure times to be dependent on observed explanatory variables. These explanatory variables can be either qualitative or quantitative. Two common regression models are the exponential regression model and the weibull regression model. Weibull regression models will be constructed and analyzed for paygrades E-8 and E-9. The primary reference for this chapter is Chapter 2 and Chapter 4 of Reference 1.

B. WEIBULL REGRESSION MODEL

1. Weibull Distribution

The Weibull distribution is a generalization of the exponential distribution that allows for a power dependence, p, of the hazard of time. The Weibull distribution is in fact an exponential distribution when $p=1$. The probability density function (pdf) of the Weibull is

$$f(t)=\lambda p(\lambda t)^{p-1}e^{(-(\lambda t)^{p})} \quad for\ t\geq 0. \qquad (4.1)$$

The hazard function is

$$\lambda(t) = \lambda p(\lambda t)^{p-1} \qquad (4.2)$$

and the survivor function is

$$S(t) = e^{(-(\lambda t)^{p})} . \qquad (4.3)$$

A convenient check to verify that data fits the Weibull distribution is to plot the log[-log $\hat{S}(t)$] versus log t($\hat{S}(t)$ is the empirical estimate of the survivor function). If the data can be fitted by the Weibull distribution, the plot should provide a relatively straight line with the slope of the line roughly estimating the value of p.

2. Log-Linear Regression Model

The Weibull distribution can be generalized to obtain a Weibull regression model by allowing the failure rate to be a function of a vector of qualitative personnel characteristics, **X**. In this application **X** is a row vector consisting of the binary variables $X_{aviation}$, $X_{administrative}$, $X_{technical}$, $X_{skilled}$, and $X_{ogprofessional}$ which take on a value of 1 if the observation falls into the subscripted category, and are 0 otherwise.

If the conditional (on **X**) hazard is

$$\lambda(t|X) = \lambda p(\lambda t)^{p-1} e^{X\beta}$$

then the conditional density of T is

$$f(t|X) = \lambda p(\lambda t)^{p-1} e^{X\beta} e^{-[(\lambda t)^p e^{X\beta}]} \qquad (4.4)$$

where $\beta = (\beta_1, \ldots, \beta_n)$ is a vector of regression parameters. In terms of Y=log T, equation (4.4) is the linear model

$$Y = \alpha + X\beta^* + \sigma W \qquad (4.5)$$

where $\alpha = -\log \lambda$, $\sigma = p^{-1}$, $\beta^* = -\sigma\beta$, and W has the extreme value distribution. In equation (4.5) the personnel characteristics act additively on the time scale, vice multiplicatively in equation (4.4). [Ref. 1]

3. Modelled Survivor Function

Equation (4.5) is known as an accelerated failure time model. It is characterized by the scaling of the failure time

$$T = e^{-X\beta} T_0$$

where T_0 is a failure time from a baseline when the values of X are all zero. Using this time scaling the modelled survivor function is

$$S(t) = e^{-(\lambda e^{-\beta X} t)^{p}} \qquad (4.6).$$

In equation (4.6), S(t) can be estimated by inserting appropriate values for **X** and t, and estimated values of β, λ, and p. This will produce a modelled survivor function which can be plotted to produce curves like those in Chapter III. If this modelled survivor function is similar (in shape and magnitude) to the empirical survivor function, then the model can be used. If it is not similar, then it possibly should not be used.

C. RESULTS

The partitioned data sets were regressed using the SAS LIFEREG procedure [Ref. 4] to obtain estimates for β, λ, and p. These values along with appropriate values for **X** were inserted into equation (4.6) to estimate modelled survivor functions. Figures 8 through 13 on pages 43 through 46 contain plots of these modelled functions against empirical survivor functions. Generally the modelled survivor functions do not fit the data as well as desired.

Nonetheless, modelled survivor functions for these two paygrades will be used pending the development of sharper alternatives. The partitioned data sets are so small, it is felt that a model will produce the most reliable and stable probability estimates.

V. BINOMIAL COUNTING MODEL

A. INTRODUCTION

The Binomial Distribution provides a means for calculating expected values and variances of sums of random variables that have one of two outcomes: success or failure. This study is concerned with estimating the number of enlisted who separate on any given month (or conversely the number who do not separate). This chapter develops a simple model, based on principals of the Binomial Distribution, that calculates monthly survival figures per paygrade, and provides an example of how the model works. The primary reference is Chapter 3 of Reference 5.

B. BINOMIAL DISTRIBUTION

A binomial experiment is an experiment that possesses the following properties:

- the experiment consists of n identical trials,
- each trial results in one of two outcomes, a success or failure,
- the probability of success on a single trial is equal to p and remains the same from trial to trial,

- trials are independent, and

- the random variable of interest is *Y*, the number of successes in *n* trials. The expected value of Y is E(Y)=np, where n is the number of trials and p is their corresponding probability of success. [Ref. 5]

A person separating from the CG meets these standards and can be considered a Bernoulli trial. Each term *trial* results in one of two outcomes. A *success* is marked if the person does not separate, a *failure* when the person leaves the CG. The product-limit-estimate of the survivor function provides the probability of a person surviving to and beyond a time, *t*; $P(T \geq t)$ for $t=1,\ldots\infty$. This *probability* estimate remains the same for each trial (ie. for all persons with length of services equal to time *t*). The decision by a person to separate (or the CGs decision to separate a person) is not affected by another persons choice and thus *independent* of any other persons decision. And finally, *Y* equals the number of persons not separating within a paygrade who have survived beyond a time *t*. Using these definitions a basic model can be constructed that generates an expected value and variance for the number of persons that do not separate on a given month.

C. COUNTING MODEL

Paygrades are partitioned in two steps to form data cells that consist of independent and identical trials required of the Binomial Distribution. The first step separates each paygrade into the single personnel characteristic examined in the survival analysis. These will be called sets. The second step further partitions the paygrade by the amount of time each person has served, or Time in Service (TIS) cells. Figure 14 provides a tree diagram example of this partitioning scheme.

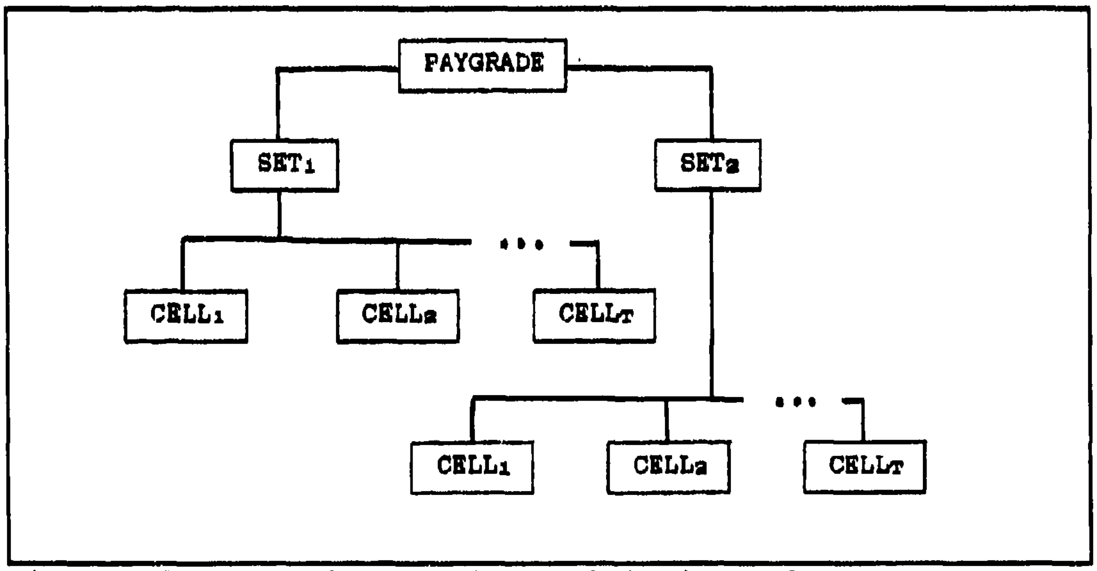

Figure 14. Counting Model Partitioning Scheme

T is the maximum TIS any one person has within each set. Paygrades E-1 and E-2 are partitioned by race to form two data sets each; paygrade E-3 is partitioned by MOS obtainment; and paygrades E-4 through E-9 are partitioned by MOS group. Each cell has a corresponding probability of success, p_j, (either estimated from the actual data or generated by a model) which was defined in Section 4.D.1 to be the probability of survival being greater than or equal to t. Given n_j, the expected number of persons who fall into cell j, the expected number of successes (persons that do not separate) in each cell is

$$E(Y_j) = n_j p_j \quad ; \quad j=1,\ldots,T. \qquad (4.7)$$

The cell variance is

$$Var(Y_j) = n_j p_j q_j \qquad (4.8)$$

where $q_j = 1 - p_j$. The expected number of successes within each set is the sum of the cell expected values

$$E(Y_{set}) = \sum_{j=1}^{T} E(Y_j) \qquad (4.9)$$

and by summing cell variances the set variance is

$$Var(Y_{s\alpha}) = \sum_{j=1}^{T} Var(Y_j) \qquad (4.10)$$

The expected number of successes per paygrade is calculated by summing set expected values. Paygrade variance is the sum of all set variances.

Appendices C and D contain procedures for generating product limit estimates which correspond to the p_{ij}s.

D. EXAMPLE

Data from paygrade E-1 is used to illustrate this counting model. Paygrade E-1 contains 2 paygrade sets. Set_1 contains records for caucasians and Set_2 contains records for non-caucasians. Table 6 and Table 7 on page 32 contain frequency counts of the number of persons falling within each cell per set.

TABLE 6. PAYGRADE CELL FREQUENCY COUNT CAUCASIAN E-1

Month	Freq	Month	Freq	Month	Freq	Month	Freq
1	53	11	2	27	1	39	3
2	339	12	3	30	2	40	2
3	148	15	3	31	1	42	2
4	8	16	2	32	1	43	1
5	2	21	4	34	1	48	1
6	1	23	3	36	1	60	1
7	1	25	2	38	1	>60	5

TABLE 7. PAYGRADE CELL FREQUENCY COUNT NON-CAUCASIAN E-1

Month	Freq	Month	Freq	Month	Freq	Month	Freq
1	2	4	6	16	1	52	1
2	87	13	1	18	1	>52	0
3	35	14	1	20	1		

Tables 8 and Table 9 on page 33 contain the estimated p_j values for each data sets.

TABLE 8. PRODUCT LIMIT ESTIMATES PAYGRADE E-1 CAUCASIAN

Month	P	Month	P	Month	P
1	.88	15	.36	27	.256
2	.645	16	.345	28	.243
3	.5	17	.337	29	.23
4	.468	18	.33	30	.225
5	.441	19	.326	31	.22
6	.431	20	.318	32	.21
7	.424	21	.314	33	.2
8	.417	22	.294	35	.2
11	.417	23	.294	36	.1
12	.393	24	.286	48	.1
13	.389	25	.265	61	.05
14	.375	26	.256	62	0

TABLE 9. PRODUCT LIMIT ESTIMATES PAYGRADE E-1 NON-CAUCASIAN

Month	P	Month	P	Month	P
1	.889	13	.225	37	.107
2	.646	14	.213	39	.107
3	.444	15	.173	40	.089
4	.403	16	.173	43	.089
5	.368	17	.173	44	.071
6	.344	18	.159	48	.071
7	.344	19	.159	49	.054
8	.332	20	.143	64	.054
10	.332	30	.143	65	.027
11	.249	31	.125	66	0
12	.249	36	.125		

Substituting $\hat{n}_j$ the values from tables 6 and 7 for the n_j values and $\hat{p}_j$ the estimated probability values from Tables 8 and 9 into equation (4.9), the estimated expected value of each set is

$$\hat{E}(Y_1)=\sum_{j=1}^{60}\hat{n}_j\hat{p}_j=353.475$$

$$\hat{E}(Y_2)=\sum_{j=1}^{52}\hat{n}_j\hat{p}_j=76.905$$

where set_1 corresponds to caucasians and set_2 corresponds to non-caucasians. Using equation (4.10) the $\hat{V}ar(Y_1)$ = 129.5 and $\hat{V}ar(Y_2)$ = 30.97. The overall paygrade estimated expected value and variance is the sum of the set values

$$\hat{E}(Y_{E-1})=\sum_{i=1}^{2}\hat{E}(Y_i)=430.38$$

$$V\hat{a}r(Y_{E-1})=\sum_{i=1}^{2}V\hat{a}r(Y_i)=160.47.$$

For working purposes these values are rounded down.

VI. CONCLUSIONS AND RECOMMENDATIONS

A. CONCLUSIONS

There are approximately 30,100 enlisted persons in the CG at any given time. A method is needed to predict monthly attrition rates to maintain the CG enlisted hierarchical structure. Survival analysis was used to examine the characteristics of sex, marital status, race, and MOS group behavior of CG enlisted personnel. The results were fairly interesting. Males and married persons tended to have higher probabilities of survival than did females and individuals who were not married. Non-caucasians in most of the paygrades had higher survival probabilities than did caucasians. MOS groups demonstrated different characteristic behaviors among different paygrades. MOS groups in paygrades E-4 and E-5 behaved about the same up to 48 months. After 48 months all groups showed a rapid decline in survivor probabilities with the aviation group having the highest probabilities and the technical group having the lowest. This drop is indicative of the end of a four year enlistment when individuals can usually separate from the CG. MOS groups in paygrades E-6 through E-9 all had very high probabilities of survival and demonstrated a decline at certain months. E-6 showed this decline at about 200 months and E-7 through E-9 demonstrated this at about 240 months. The drop in probabilities at 240 months is easily explained. Many Coast Guardsmen retire at 20 years and receive a substantial pension and many important benefits (ie. medical coverage for the person and his or her dependents). The drop at 200 months in paygrade E-7 is an anomoly. An E-7 at this point in his/her career has almost 17 years of service time and has only about three more years to serve before retirement. Separating at 17 years seems to be financially inappropriate because the person will not receive a pension nor benefits if they separate.

Modelled survivor functions were developed for paygrades E-8 and E-9 because these paygrades possessed small population sizes. Despite the fact that the models produced generally poor fits, it was felt that these survivor functions would be more reliable than survivor functions generated from very small data sizes. Therefore models were used to calculate attrition in paygrades E-8 and E-9.

The binomial counting model provides an effective tool for calculating monthly attrition figures. Expected values and variances can be calculated by partitioning each paygrade into several binomial experiments. This model is easy to use and can be expanded in future studies.

B. RECOMMENDATIONS

As a first study the goals of this thesis were met. However further studies are required to better analyze CG enlisted attrition. The counting model could be adjusted or refined based on findings from the recommendations listed in this section.

1. Recommendation 1

The counting model needs to be verified. Personnel data from early months of FY 92 need to be processed through the model to verify that it provides better results than the current calculation method.

2. Recommendation 2

Additional historical data needs to be analyzed to provide temporal stability to the study. The counting model produced in this study assumes that the characteristic behaviors of individuals identified in FY 91 will repeat themselves in FY 92.

3. Recommendation 3

A more manageable data base needs to be identified that allows for accessible man-machine interface. An accessible data base allows the analyst to verify, manipulate, and easily examine data entries as an on-going task. This provides data reliability. The CG PPC is not responsible for maintaining this type of data base, thus PWP should examine this issue.

4. Recommendation 4

Individual MOSs within MOS groups need to be examined to ensure they demonstrate similar attrition behaviors.

5. **Recommendation 5**

Combined personnel characteristics need to be analyzed for survival behaviors and seasonal attrition trends need to be examined.

6. **Recommendation 6**

The data base needs to be extended in time. At present there is no justification for assuming time homogeneity of rates. Such an extension would also allow the study of some important exogenous variables, such as unemployment rate and inflation.

7. **Recommendation 7**

The educational levels of Coast Guardsmen should be analyzed for their effect on attrition behaviors. High school graduates, alternate high school degree holders, and non-high school graduates need to be considered to determine if there are differences in attrition behaviors amongst these groups.

APPENDIX A. PROCEDURES FOR ACCESSING DATA

1. Data is stored on two magnetic disks. To access disks:

   a. login to mainframe computer
   b. type "getmvs mvs215 (separation data tape)" or "getmvs mvs217" (active duty data tape)

2. To run a program that has a huge output file, ie. a file that will not fit on an A-disk:

   a. type "tdisk 6" to get a temporary disk with 6 cylinders
   b. switch A-disk with temporary disk (say temp disk was assigned "D 130"
      1. type "access 130 A"
      2. type "access 191 D"
   c. copy program from the now D-disk to the newly created A-disk
   d. run program
   e. switch disks back when finished by reversing procedure

   in 2.b

APPENDIX B. FIGURES

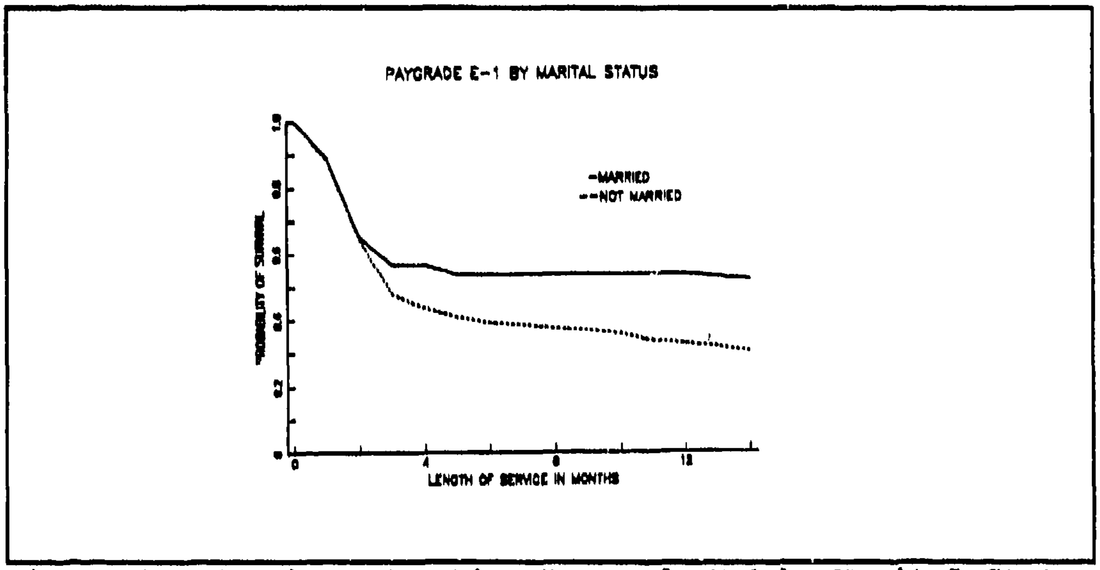

Figure 1. Survivor Function Paygrade E-1 by Marital Status

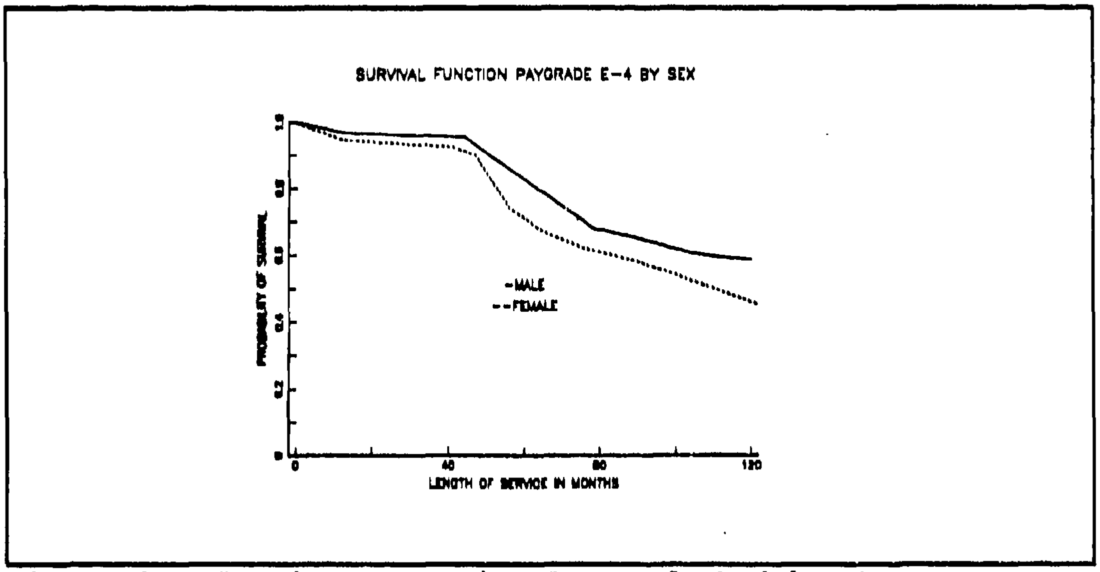

Figure 2. Survivor Function Paygrade E-4 by Sex

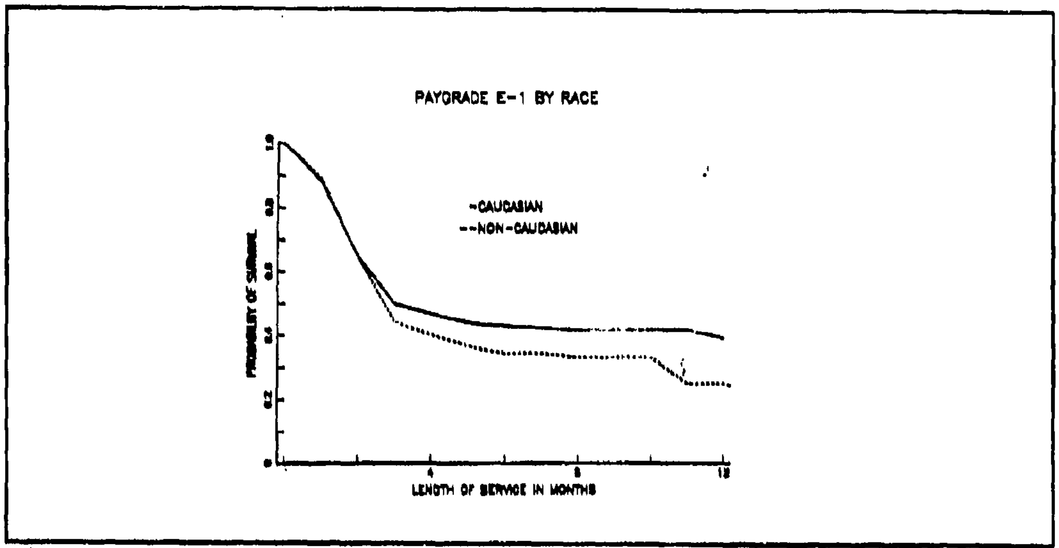

Figure 3. Survivor Function Paygrade E-1 by Race

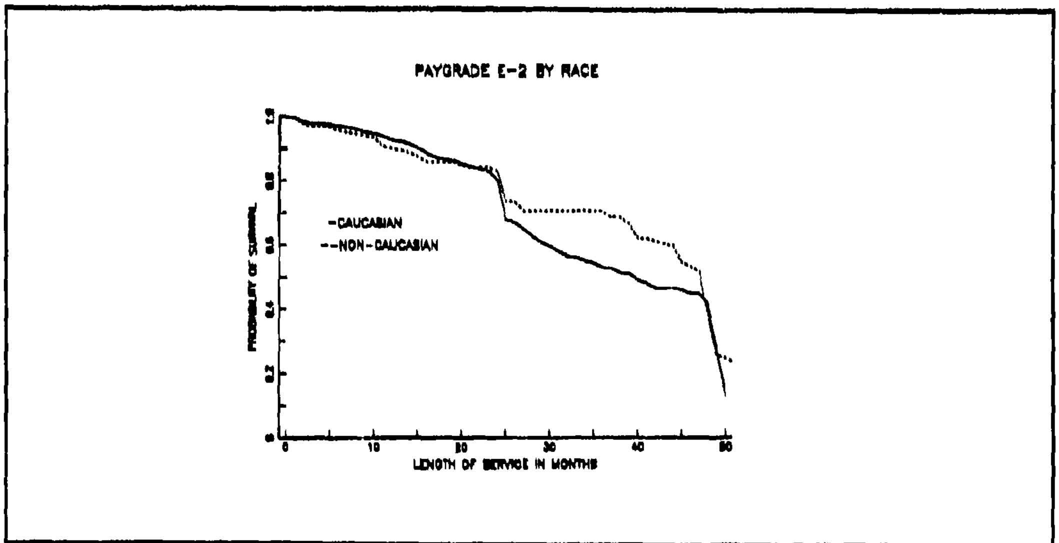

Figure 4. Survivor Function Paygrade E-2 by Race

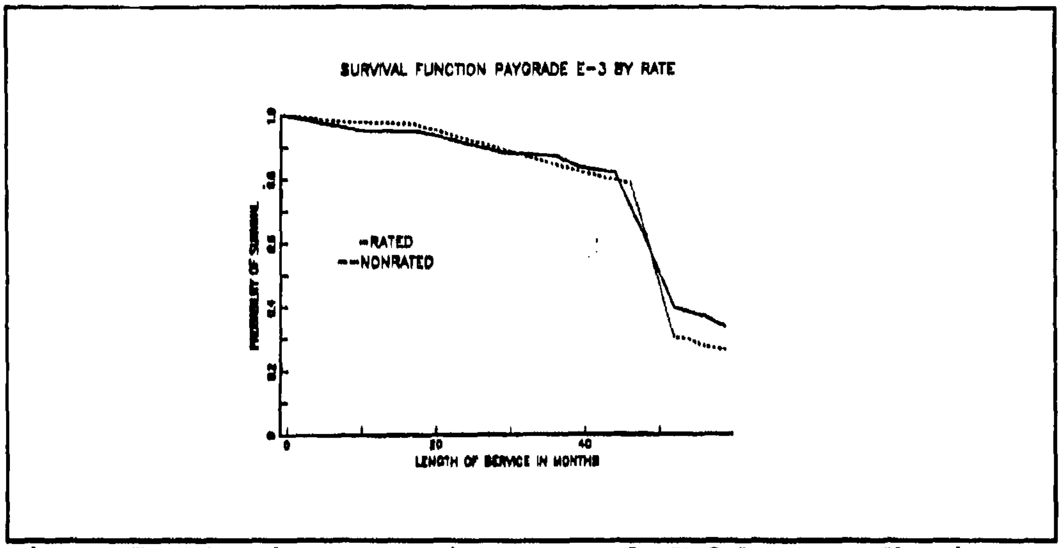

Figure 5. Survivor Function Paygrade E-3 by Rate Obtainment

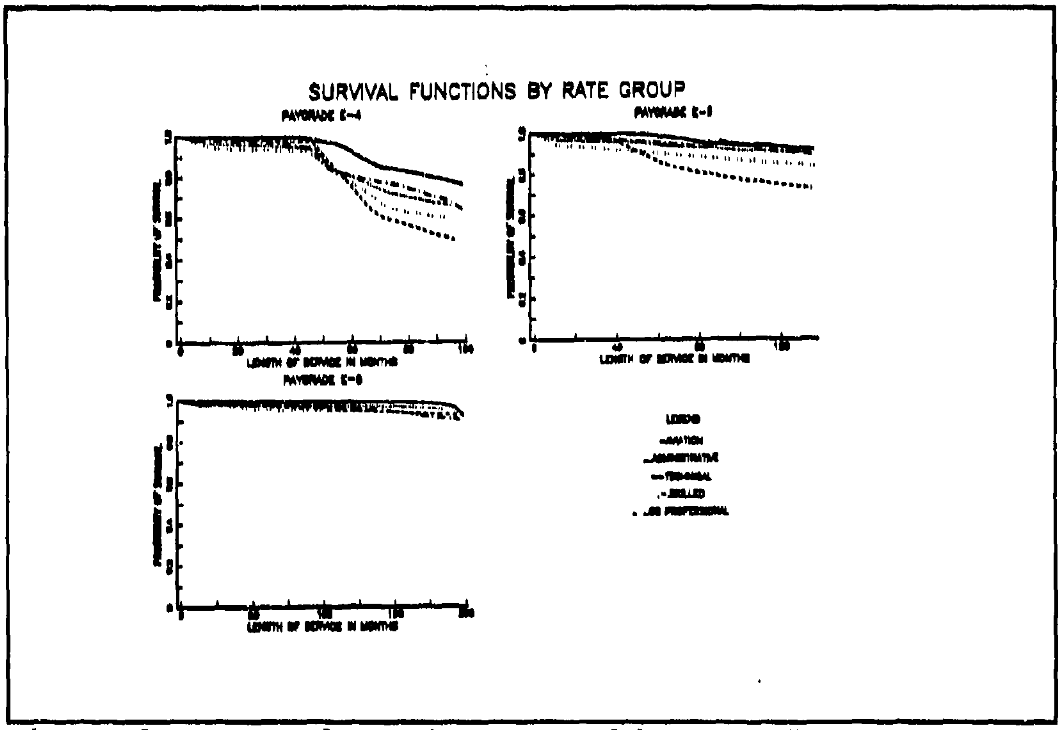

Figure 6. Paygrades E-4, E-5, E-6 by Rate Group

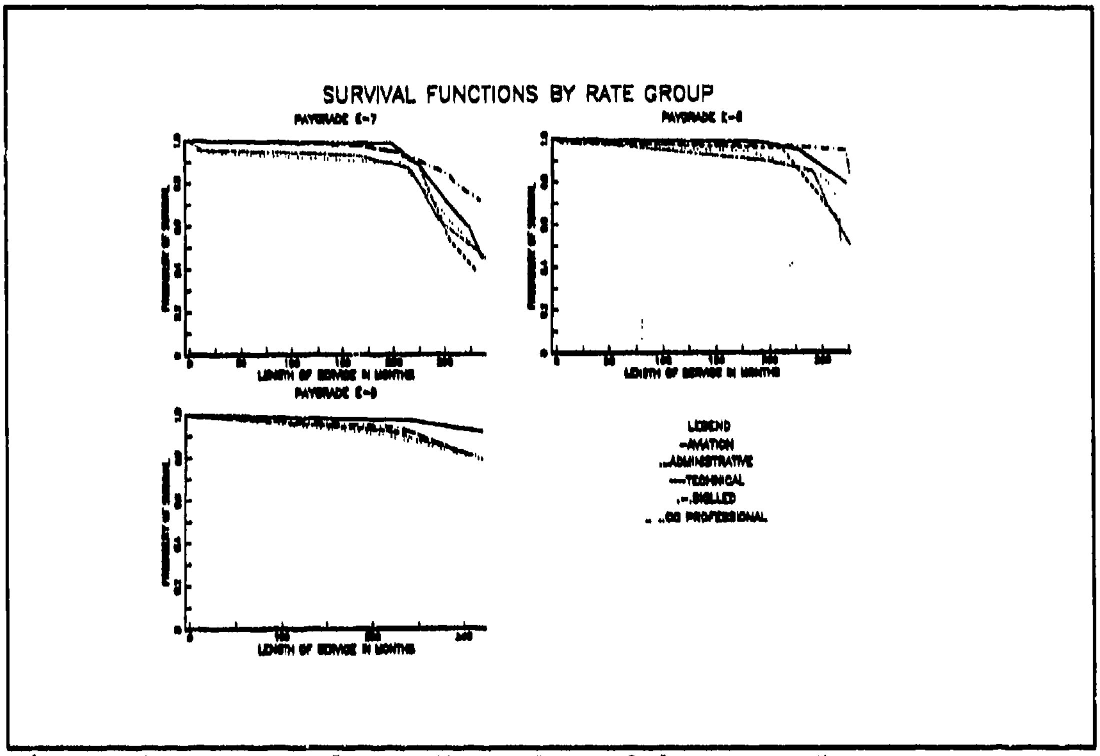

Figure 7. Paygrades E-7, E-8, E-9 by Rate Groups

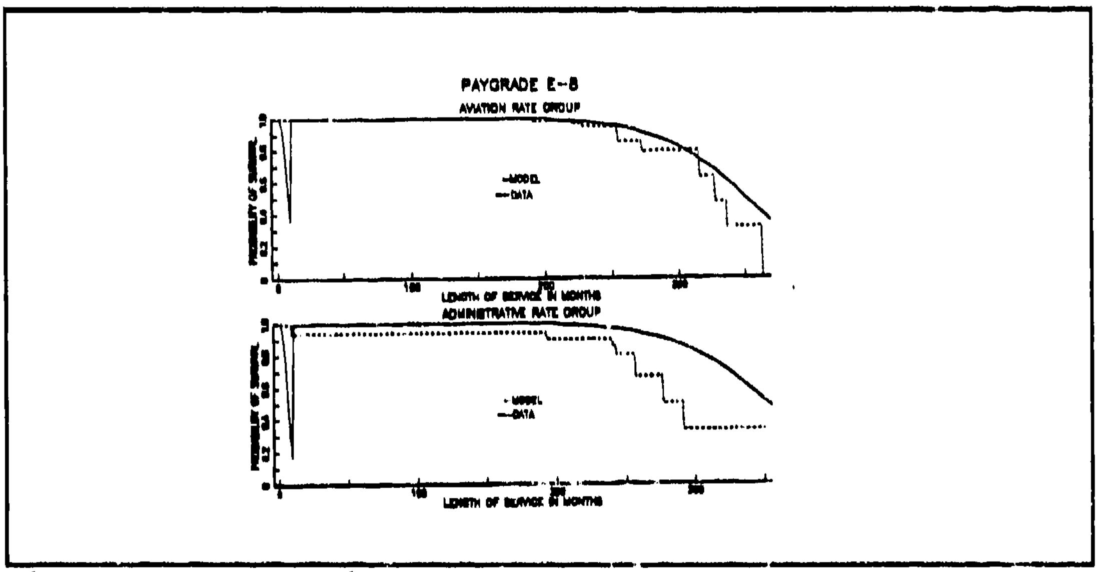

Figure 8. Model Fit Paygrade E-8

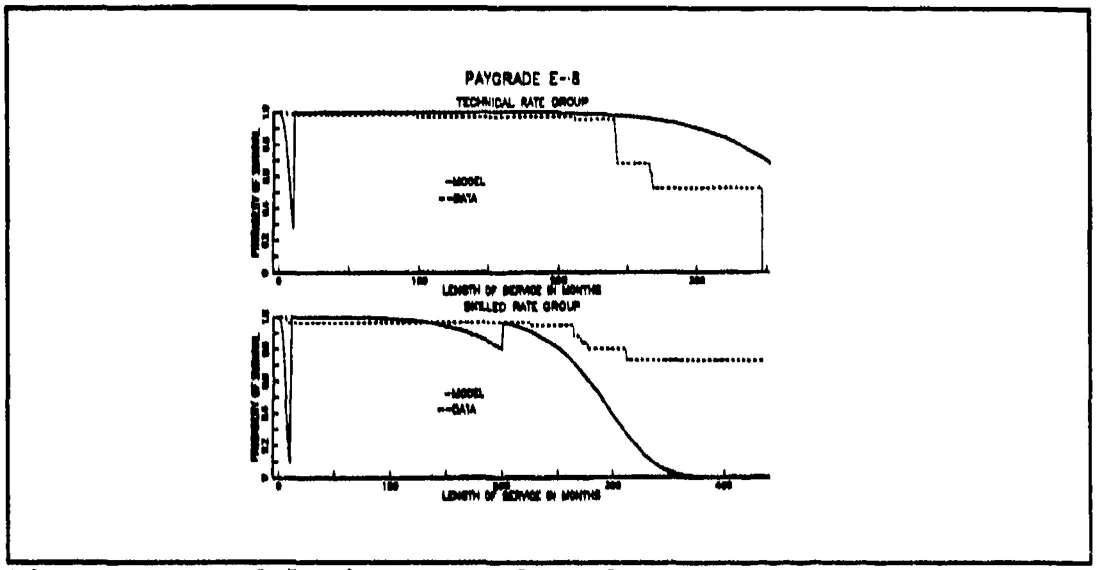

Figure 9. Model Fit Paygrade E-8

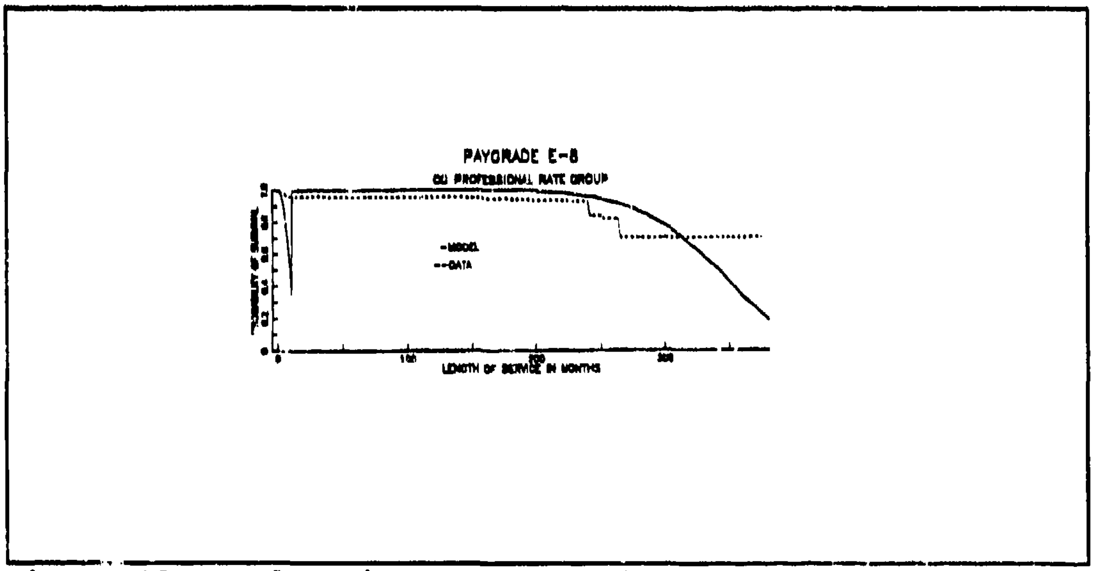

Figure 10. Model Fit Paygrade E-8

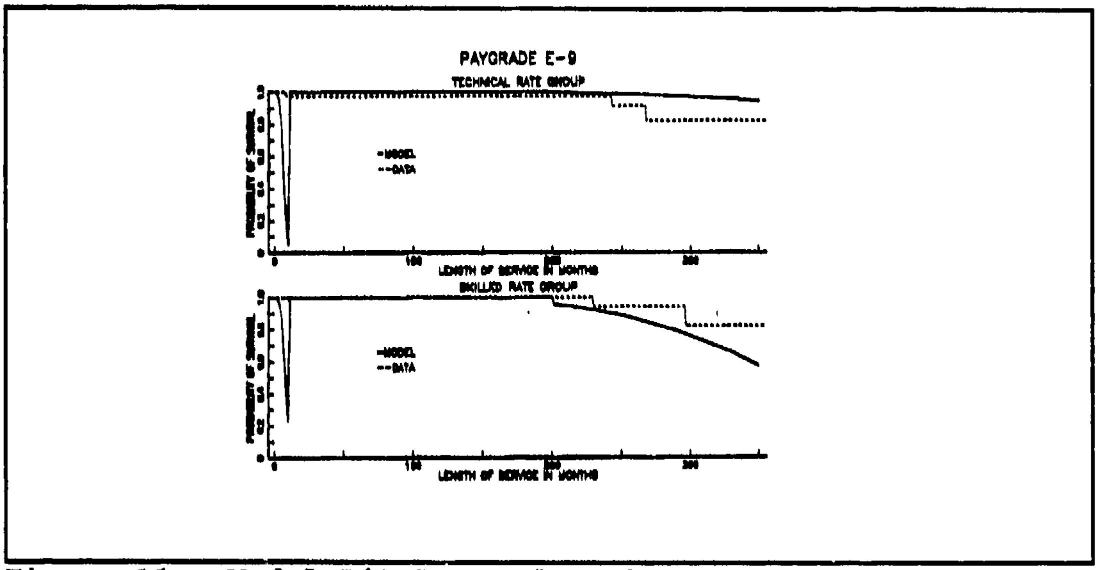

Figure 11. Model Fit Paygrade E-9

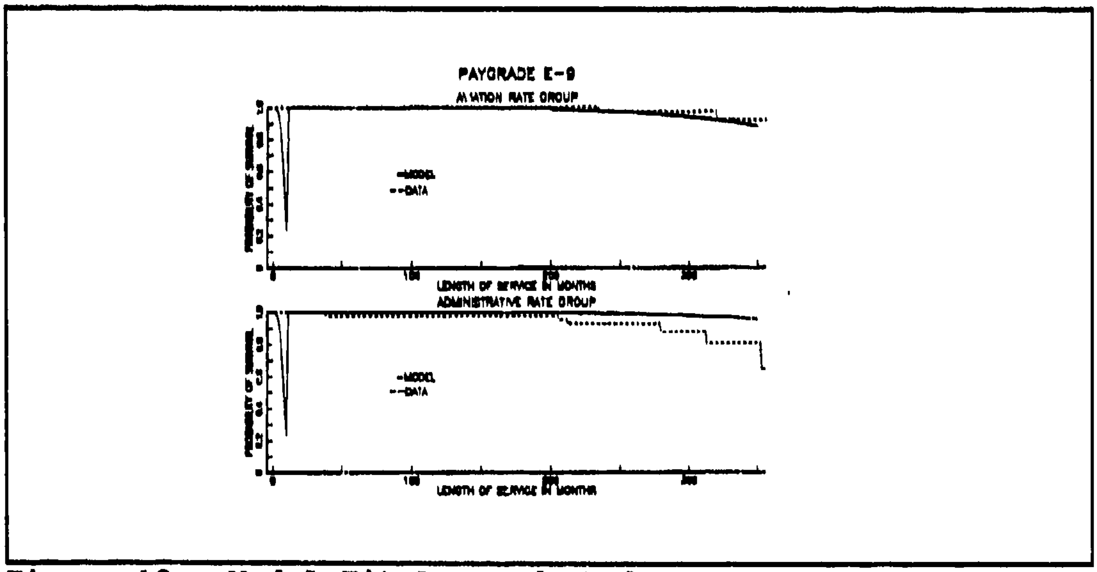

Figure 12. Model Fit Paygrade E-9

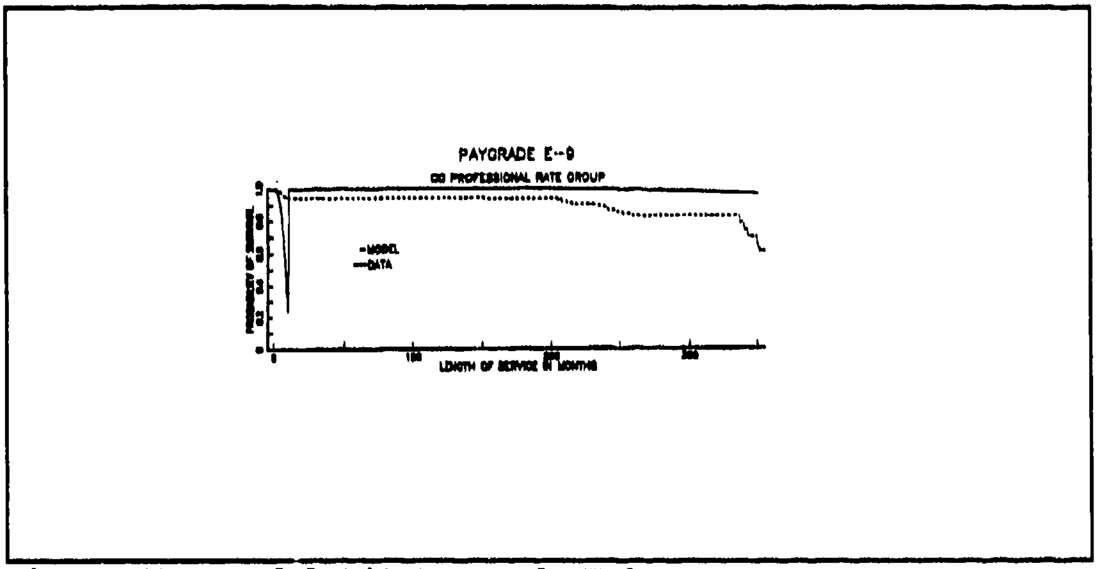

Figure 13. Model Fit Paygrade E-9

APPENDIX C. PROCEDURES FOR GENERATING SURVIVAL FUNCTIONS

```
FILE: LIFETEST CODE2    A1


TITLE'E-2 BY RACE' /* THIS SAS CODE TAKES A DATA SET AND GENERATES
SURVIVOR FUNCTIONS FOR THE DATA. THIS CODE WAS USED FOR PAYGRADE E-2
AND CREATES SURVIVAL FUNCTIONS FOR CAUCASIANS AND NON-CAUCASIANS.*/;

DATA PO2;

INPUT RATE RANK SEX AGE MAR RACE MONTHS CEN /*THIS LINE DEFINES THE
VARIABLES WITHIN THE DATA SET.*/;

GROUP=(RACE=5) /*THIS SEPARATES ALL CAUCASIANS FROM NON-CAUCASIANS,
5 INDICATES CAUCASIAN. THIS SPLITS THE DATA SET INTO 2 SET; 1 CONTIANS
CAUCASIANS, THE OTHER CONTAINS ALL OTHER MINORITITES.*/;

CARDS /*TELLS SAS THAT THE DATA SET FOLLOWS. DATA SET MUST BE PLACED
IMMEDIATELY AFTER THIS LINE AFTER THE SEMICOLON, AND MUST HAVE A
SEMICOLON PLACED IMMEDIATELY AFTER THE LAST FILE.*/ ;


;
 PROC LIFETEST PLOT=(S) MAXTIME=48 OUTSURV=PO2EST /*THIS IS THE MAIN CODE
 IT CREATES A SURVIVAL FUNCTION PLOT (S), AND LIMITS THE X-AXIS OF THE
PLOT TO 48 MONTHS. ALSO STORE THE PRODUCT LIMIT ESTIMATES IN A SAS FILE
CALLED PO2EST.*/   ;

TIME MONTHS*CEN(1) /*DEFINES THE TIME VARIABLE (MONTHS), AND SETS THE
CENSORING VARIABLE WITHIN THE DATA SET EQUAL TO 1 */  ;

STRATA GROUP /*PERFORMS THE PROCEDURE ON ALL DEFINED GROUPS. IN THIS CASE
THERE ARE ONLY 2 GROUPS*/   ;

PROC PRINT DATA=PO2EST /*PRINTS THE DATA SET PO2EST INTO A CMS FILE*/;
```

```
FILE: LIFETEST CODE     A1


TITLE'PAYGRADE E-9 BY RATE GROUP ' /*SAME AS PREVIOUS CODE EXCEPT
IT DIVIDES THE PAYGRADE INTO 5 GROUPS; 1 FOR EACH MOS GROUP. GROUP 1
IS THE AVIATION PEOPLE, 2=ADMIN ,3=TECHNICAL, 4=SKILLED,5=CO PRO */ ;

DATA PO9;
INPUT RATE RANK SEX AGE MAR RACE MONTHS CEN;
IF RATE=560|RATE=570|RATE=520|RATE=530|RATE=550 THEN GP=1;
ELSE IF RATE=360|RATE=420 THEN GP=2;
ELSE IF RATE=790|RATE=180|RATE=140|RATE=270|RATE=240|RATE=280 THEN GP=3;
ELSE IF RATE=500|RATE=840|RATE=870 THEN GP=4;
ELSE GP=5;
CARDS;
;
 PROC LIFETEST PLOT=(S) MAXTIME=300  OUTSURV=PO9EST;
TIME MONTHS*CEN(1);
STRATA GP;
PROC PRINT DATA=PO9EST;
```

```
      PROGRAM MODFIT2

C     THIS PROGRAM GENERATES MODELLED PRODUCT LIMIT ESTIMATES FOR
C     PAYGRADES WHICH HAVE BEEN PARTITIONED BY 5 MOS GROUPS: AVIATION,
C     ADMINISTRATIVE, TECHNICAL, SKILLED, CO PROFESSIONAL.  IT USES AS
C     INPUT REGRESSION COEFFICIENTS (BETA1 THROUGH BETA13), SHAPE
C     PARAMETERS (P1 AND P2), SCALE PARAMETERS (LAMBDA1 AND LAMBDA2),
C     AND A TIME PARAMETER (T). THIS PROGRAM IS ALSO DESIGNED TO WORK
C     ON DATA THAT FIT A 3 PIECE WEIBULL.  VARIABLES BETA1 THRU BETA4,
C     P1, AND LAMBDA1 ARE VALUES FOR THE FIRST WEIBULL PIECE (P IS THE
C     TIME VALUE OF WHERE THE FIRST BREAK OCCURS). VARIABLES BETA5 THRU
C     BETA8, P2, AND LAMBDA2 (AND N) CORRESPOND TO THE SECOND PIECE.
C     VARIABLES BETA9 THRU BETA12, P3, LAMBDA3 (AND N) CORRESPOND TO THE
C     LAST WEIBULL PIECE.  VARIABLE BETA13 EQUALS ZERO AND CORRESPONDS
C     TO THE CO PROFESSIONAL COEFFICIENT.  THE VARIABLE SHAT IS THE
C     MODELLED PRODUCT LIMIT ESTIMATE: SHAT1 FOR THE AVIATION MOS
C     GROUP, SHAT2 FOR ADMIN GROUP, SHAT3 FOR TECHNICAL GROUP, SHAT4 FOR
C     THE SKILLED GROUP, AND SHAT5 FOR THE CO PRO GROUP.  SHAT VALUES AR
C     CALCULATED FOR TIME PERIODST, (T=1,....,N), WHERE N IA THE LAST
C     TIME VALUE NEEDED FOR THE PAYGRADE.
C     THE ESTIMATES ARE WRITTEN ON THE CMS FILES: FT88F001, FT89F001,
C     AND FT90F001.

      INTEGER I,M,N,P

      REAL LAMBDA1,P1,BETA1,BETA2,BETA3,BETA4,BETA5,BETA13
      REAL LAMBDA2,P2,BETA6,BETA7,BETA8,BETA9,BETA10,BETA11,BETA12
      REAL SHAT1,SHAT2,SHAT3,SHAT4,SHAT5,LAMBDA3,P3

C     FOLLOWING ESTABLISHES ALL VARIABLE VALUES.

      PARAMETER(M=10,LAMBDA1=.1007 ,P1=2.685)
      PARAMETER(BETA1=0.0,BETA2=-.1930,BETA3=-.0833,BETA4=-.3102)
      PARAMETER(N=200 ,LAMBDA2=0.00003 ,P2=4.635)
      PARAMETER(BETA5=-.0545,BETA6=.119,BETA7=-.0519,BETA8=-4.7896)
      PARAMETER(P=420,LAMBDA3=.0029 ,P3=0.0336 )
      PARAMETER(BETA9=.0254,BETA10=.0397,BETA11=.106,BETA12=-.1665)
      PARAMETER(BETA13=0.0    )

C     CALCULATES SHATS FOR FIRST WEIBULL PIECE.

      DO 10 I=1,M
         SHAT1=EXP(-((LAMBDA1*I*EXP(-BETA1))**P1))
         SHAT2=EXP(-((LAMBDA1*I*EXP(-BETA2))**P1))
         SHAT3=EXP(-((LAMBDA1*I*EXP(-BETA3))**P1))
         SHAT4=EXP(-((LAMBDA1*I*EXP(-BETA4))**P1))
         SHAT5=EXP(-((LAMBDA1*I*EXP(-BETA13))**P1))
       WRITE(88,91)I,SHAT1,SHAT2,SHAT3,SHAT4,SHAT5
   10 CONTINUE

C     CALCULATES SHATS FOR SECOND WEIBULL PIECE.

      DO 20 I=M+1,N
         SHAT1=EXP(-((LAMBDA2*I*EXP(-BETA5))**P2))
         SHAT2=EXP(-((LAMBDA2*I*EXP(-BETA6))**P2))
         SHAT3=EXP(-((LAMBDA2*I*EXP(-BETA7))**P2))
         SHAT4=EXP(-((LAMBDA2*I*EXP(-BETA8))**P2))
         SHAT5=EXP(-((LAMBDA2*I*EXP(-BETA13))**P2))
       WRITE(89,91)I,SHAT1,SHAT2,SHAT3,SHAT4,SHAT5
   20 CONTINUE

C     CALCULATES SHATS FOR LAST PIECE.

      DO 30 I=N+1,P
         SHAT1=EXP(-((LAMBDA3*I*EXP(-BETA9))**P3))
         SHAT2=EXP(-((LAMBDA3*I*EXP(-BETA10))**P3))
         SHAT3=EXP(-((LAMBDA3*I*EXP(-BETA11))**P3))
         SHAT4=EXP(-((LAMBDA3*I*EXP(-BETA12))**P3))
         SHAT5=EXP(-((LAMBDA3*I*EXP(-BETA13))**P3))
       WRITE(90,91)I,SHAT1,SHAT2,SHAT3,SHAT4,SHAT5
   30 CONTINUE
   91 FORMAT(1X,I3,3X,F6.4,3X,F6.4,3X,F6.4,3X,F6.4,3X,F6.4)
```

APPENDIX D. PROCEDURES FOR CONDUCTING REGRESSION ANALYSIS

```
      PROGRAM REGPREP

C     THIS CODE SEPARATES A PAYGRADE DATA SET BY MOS GROUP AND CREATES
C     DATA SETS NEEDED TO CONDUCT THE SAS LIFEREG PROCEDURE. AGAIN THIS
C     CODE IS DESIGNED FOR A 3 PIECE WEIBULL. THE INPUT FILE (#2)
C     CONTAINS THE PAYGRADE DATA SET.  3 OUTPUT DATA SETS ARE CREATED TO
C     CONFORM TO THE 3 PIECE WEIBULL. THE VARIABLE N IS THE NUMBER OF
C     FILES IN THE PAYGRADE.

      INTEGER I,N,RATE,MONTHS,CEN
      PARAMETER(N=448)
      OPEN(2,FILE='POS DATA A')
      DO 10 I=1,N
         READ(2,92) RATE,MONTHS,CEN

C     THIS IS THE FIRST PIECE OF THE WEIBULL

         IF (MONTHS .LE. 10) THEN
           IF(RATE .EQ. 520 .OR. RATE .EQ. 560 .OR. RATE .EQ.570 .OR.
     1        RATE .EQ. 530 .OR. RATE .EQ. 550)THEN
              WRITE(26,93) MONTHS,CEN
           ELSEIF(RATE .EQ. 360 .OR. RATE .EQ. 420) THEN
              WRITE(26,94) MONTHS,CEN
           ELSEIF(RATE .EQ. 270 .OR. RATE .EQ. 240 .OR. RATE .EQ. 180
     1          .OR. RATE .EQ. 790 .OR. RATE .EQ. 140 .OR. RATE .EQ.
     1          280) THEN
              WRITE(26,95) MONTHS,CEN
        ELSEIF(RATE .EQ. 870 .OR. RATE .EQ. 500 .OR. RATE .EQ. 840)THEN
              WRITE(26,96) MONTHS,CEN
           ELSE
            WRITE(26,97) MONTHS,CEN
           ENDIF

C     THIS IS THE SECOND PIECE OF THE WEIBULL.

         ELSEIF(MONTHS .GT. 10 .AND. MONTHS .LE. 200)THEN
           IF(RATE .EQ. 520 .OR. RATE .EQ. 560 .OR. RATE .EQ.570 .OR.
     1        RATE .EQ. 530 .OR. RATE .EQ. 550)THEN
              WRITE(27,93) MONTHS,CEN
           ELSEIF(RATE .EQ. 360 .OR. RATE .EQ. 420) THEN
              WRITE(27,94) MONTHS,CEN
           ELSEIF(RATE .EQ. 270 .OR. RATE .EQ. 240 .OR. RATE .EQ. 180
     1          .OR. RATE .EQ. 790 .OR. RATE .EQ. 140 .OR. RATE .EQ.
     1          280) THEN
              WRITE(27,95) MONTHS,CEN
        ELSEIF(RATE .EQ. 870 .OR. RATE .EQ. 500 .OR. RATE .EQ. 840)THEN
              WRITE(27,96) MONTHS,CEN
           ELSE
            WRITE(27,97) MONTHS,CEN
           ENDIF

C     THIS IS THE LAST PIECE OF THE WEIBULL.

         ELSE
           IF(RATE .EQ. 520 .OR. RATE .EQ. 560 .OR. RATE .EQ.570 .OR.
     1        RATE .EQ. 530 .OR. RATE .EQ. 550)THEN
              WRITE(28,93) MONTHS,CEN
           ELSEIF(RATE .EQ. 360 .OR. RATE .EQ. 420) THEN
              WRITE(28,94) MONTHS,CEN
           ELSEIF(RATE .EQ. 270 .OR. RATE .EQ. 240 .OR. RATE .EQ. 180
     1          .OR. RATE .EQ. 790 .OR. RATE .EQ. 140 .OR. RATE .EQ.
     1          280) THEN
              WRITE(28,95) MONTHS,CEN
        ELSEIF(RATE .EQ. 870 .OR. RATE .EQ. 500 .OR. RATE .EQ. 840)THEN
              WRITE(28,96) MONTHS,CEN
           ELSE
            WRITE(28,97) MONTHS,CEN
           ENDIF
         ENDIF
   10 CONTINUE

C     THIS IS THE FORMAT FOR THE READ/INPUT STATEMENTS.

   92 FORMAT(1X,I3,13X,I3,1X,I1)

C     THESE FORMAT STATEMENTS CREATE REGRESSION DATA SETS. MONTHS IS THE
C     DEPENDENT VARIABLE AND IS LOCATED IN THE FIRST COLUMN. THE
C     PRECEDING COLUMNS CONTAIN A 0 OR 1 IF THE PERSON FALLS INTO ONE
C     OF THE MOS GROUPS. (THE 5TH MOS GROUP IS ASSUMED AND DOES NOT SHO
C     IN THE DATA SET.

   93 FORMAT(1X,I3,1X,'1',1X,'0',1X,'0',1X,'0',1X,I1)
   94 FORMAT(1X,I3,1X,'0',1X,'1',1X,'0',1X,'0',1X,I1)
   95 FORMAT(1X,I3,1X,'0',1X,'0',1X,'1',1X,'0',1X,I1)
   96 FORMAT(1X,I3,1X,'0',1X,'0',1X,'0',1X,'1',1X,I1)
   97 FORMAT(1X,I3,1X,'0',1X,'0',1X,'0',1X,'0',1X,I1)
      END
```

```
FILE: LIFEREG  CODE     A1

 TITLE'REGRESSION OUTPUT PAYGRADE E-6' /*SAS CODE PERFORMS A REGRESSION
 ON CENSORED DATA. GENERATES REGRESSION COEFFICIENTS, SHAPE AND SCALE
PARAMETERS. OUTPUT IS VERY SMALL. */ ;

 DATA POS;
 INPUT MONTHS AVIATION ADMIN TECH SKILLED CEN;
 CARDS;
 ;
 PROC LIFEREG OUTEST=MODEL;

 CLASS AVIATION ADMIN TECH SKILLED /*SETS THE VARIABLES AVIATION,ADMIN,
 TECH, AND SKILLED AS CATEGORICAL (0 OR 1) VARIABLES */ ;

 MODEL MONTHS*CEN(1)=AVIATION ADMIN TECH SKILLED/DIST=WEIBULL /*SETS THE
 DEPENDENT VARIABLE EQUAL TO MONTH, SETS THE CENSORING VARIABLE,CEN=1,
 SETS THE INDEPENDENT VARIABLES EQUAL TO AVIATION, ADMIN, TECH, SKILLED,
 ESTABLISHES THE WEIBULL AS THE PARAMETRIC DISTRIBUTION */  ;
```

LIST OF REFERENCES

1. Kalbfleisch, J. D., and Prentice, R. L., *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, Inc., 1980.

2. Schatzoff, Martin, and Lane, Thomas P., 1986, *Reliability Analysis in GRAFSTAT*, Research Report RC 11655 (No. 52654), IBM T.J. Watson Research Center, Yorktown Heights, N.Y.

3. Hawes, Eric A., *An Application of Survival Analysis Methods to the Study of Marine Enlisted Attrition*, Master's Thesis, Naval Postgraduate School, Monterey, California, March 1990.

4. SAS Institute, Inc., *SAS User's Guide: Statistics*, Version 5 Edition, 1985.

5. Mendahall, William, and Wackerly, Dennis D., and Scheaffer, Richard L., *Mathematical Statistics with Applications*, Fourth Edition, PWS-KENT Publishing Company, 1990.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 52 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Commandant (PWP) United States Coast Guard 2100 Second St. SW Washington D.C., 20593	1
4.	Professor R.R. Read (Code OR/RE) Operations Analysis Dept. Naval Postgraduate School Monterey, CA 93943	1
5.	Professor Lyn R. Whitaker (Code OR/WH) Operations Analysis Dept. Naval Postgraduate School Monterey, CA 93943	1
6.	LT Douglas A. Blakemore COMDT (G-PWP-1) U.S. Coast Guard 2100 Second St. SW Washington, D.C., 20593	2