If you have issues viewing or accessing this file, please contact us at NCJRS.gov.

164098

U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

National Institute of Justice

Working Paper

Risk Classification of Probationers: Development of a Statistical Model Based on Management Information System Data

NIJ Working Paper 96-01

Pamela K. Lattimore
Richard L. Linster

January 1996

This research was supported by funds from the National Institute of Justice.

The National Institute of Justice produces working papers as a way to document findings of staff research projects. These papers may be reproduced eventually in other contexts such as journal articles or Institute publications. Opinions expressed in discussion papers are those of the authors and not necessarily those of the U.S. Department of Justice. Comments by readers are welcomed.

1 64098

RISK CLASSIFICATION OF PROBATIONERS: DEVELOPMENT OF A STATISTICAL MODEL BASED ON MANAGEMENT INFORMATION SYSTEM DATA

NIJ WORKING PAPER 96-01

Pamela K. Lattimore
National Institute of Justice
633 Indiana Avenue NW
Washington, DC 20531
(202) 307-2961
PAMLAT@UMDD.UMD.EDU

Richard L. Linster
National Institute of Justice
633 Indiana Avenue NW
Washington, DC 20531

January 1996

Points of view are those of the authors and do not necessarily reflect the opinion of the National Institute of Justice or the U.S. Department of Justice. Comments and recommendations are welcome and should be addressed to the first author.

ACKNOWLEDGMENTS

The authors of this report acknowledge with gratitude the assistance of their colleagues in the Bureau of Planning, Research and Statistics at the Florida Department of Corrections. Thanks are especially due to Bill Bales, Jenny Nimer and Kristine Leininger of the Bureau of Research for the expert advice and insight they provided and for their patience in answering the many questions that arose as the work progressed. Further, we wish to express our appreciation to Harry Dodd, Assistant Secretary for Programs of the Florida Department of Corrections, who has been the motivating spirit behind this project and has contributed greatly in keeping it focussed on its practical goals.

RISK CLASSIFICATION OF PROBATIONERS: DEVELOPMENT OF A STATISTICAL MODEL BASED ON MANAGEMENT INFORMATION SYSTEM DATA

ABSTRACT

This report describes the establishment of the methods and models for using Department of Corrections' management information system data to classify individuals by their likelihood of "failing" during probation. The classification is based on statistical estimation of the likelihood of individual failure (revocation and absconding) during the probation supervision period. These estimations are then used to classify individuals as high, medium, or low risk for supervision purposes. The study is being conducted jointly by researchers at the National Institute of Justice and researchers and administrators at the Florida Department of Corrections. While the data used in the research reported here are specific to Florida's probationers and policies, the methods developed here would seem appropriate to addressing the practical issues of risk assessment in other jurisdictions and program settings.

1.0 INTRODUCTION

This project is directed at the development of a classification system for adult probationers that uses data readily available from a Department of Corrections management information system. The classification is based on statistical estimation of the likelihood of individual failure during the probation supervision period. These estimates are then used to establish a triage system which classifies probationers as high, medium or low risks based on the individual factors found to be statistically associated with probation failure. The study is being conducted jointly by researchers at the National Institute of Justice (NIJ) and at the Bureau of Planning, Research and Statistics of the Florida Department of Corrections (FLDOC).

This report addresses the first phase of the project: The establishment of the methods (models) for determining individual failure probabilities, given a minimal set of variables to characterize each case. The analyses use case history records on offenders admitted to Florida probation during the four-year period January 1, 1991, through December 31, 1994, and entail identifying patterns in the management information system data that are statistically related to observed, in-program failure. While these data are specific to Florida's probationers and policies, the <u>methods</u> developed here would seem appropriate to addressing the practical issues of risk assessment in other jurisdictions and program settings.

On-going efforts involve translating these failure probability results into a risk classification system that reflects the policy of the FLDOC, given the constraints imposed by the size of the actively supervised population and the available probation officer resources.

The report is organized as follows: This first section provides a short background describing some important aspects of Florida's probation system. This is followed by a brief overview of the main results of the study. Hypothetical examples are given to illustrate some important practical implications that follow from translation of individual failure probabilities into a three-level risk classification system. Section 2 contains the technical description of the mathematical methods by which individual failure probabilities are calculated along with some explanation of why these particular methods were adopted.

Finally, Section 3 summarizes the work planned by researchers at the Florida Department of Corrections to convert the technical results into a fully automated system of probationer risk classification.

1.1. FLORIDA PROBATION

In mid-year 1994, the Florida Department of Corrections had about 90,000 probationers under active supervision.[1] These offenders had been convicted of a felony in state courts on charges running the gamut from serious traffic offenses to murder.[2] The most prominent offense categories among the 51,099 most recent probation admissions (from July 1, 1993, through June 30, 1994) were drug offenses and the property crimes of theft, forgery or fraud. Each of these two offense categories accounted for about 26% of all intakes. For the majority of these recent probationers (52.4%), the current probation sentence was their first commitment to the Florida Department of Corrections. The length of the average sentence was 2.5 years of supervision, although sentences ranged from less than one year to life. Finally, about 6% were split sentences with intake to probation following some term of incarceration.

The policy manual defines three levels of probation supervision[3]:

- Maximum--at least 2 personal and 2 collateral (family, employer, etc.) contacts per month.

---

[1]Descriptive statistics and other information in this section are taken variously from Florida Department of Corrections publications "1993-94 Annual Report: The guidebook to Corrections in Florida," "Florida's Community Supervision Offender Trends; Quarterly Report: Admissions July to September 1983-1993," and "Supervision: Probation and Parole Manual of Procedures."

[2]About 2% were convicted of a misdemeanor reduced from a felony charge.

[3]The policy directives quoted here are from the June 1986 revision to the manual "Supervision: Probation and Parole Manual of Procedures."

- Medium--at least 1 personal and 1 collateral contact per month.
- Minimum--at least 1 personal contact per month.

Nominally at least, current probationer classifications are calendar-driven. All new intakes are placed in maximum supervision for the first three months. Subsequently, cases are reclassified to medium supervision, although under specified conditions cases may be retained in maximum ("... justification must be clearly shown to retain a case in maximum supervision beyond 90 days."). Finally, again with certain exceptions "... cases in minimum supervision shall be reviewed every six months ... for early termination" and cases in minimum classification for one year " ... shall be recommended to the court ... for early termination."

The manual also states that "[m]inimum contact standards for probationers and parolees under supervision are based on the Workhour Formula with an overall officer caseload of 68 cases and adequate travel allowances. Caseloads exceeding this level will cause the number of contacts to be reduced proportionately diminishing protection to the community." The average caseload in FY 1993-94 was 116. At least in part, then, this study was motivated by the problem of how to allocate limited supervision resources in a way that could make the greatest contribution to public safety.

1.2. ASSESSING RISK OF FAILURE ON PROBATION

During any specified time period, a probationer can (1) be released from probation; (2) be revoked (or otherwise enter a "failure" status); or (3) continue in an active status. For purposes of this study three distinct official actions denoting failure were defined:

- Revocation for a new arrest,
- Revocation for a technical violation, and
- Issuance of an absconder warrant.

Each of these actions defining probation failure results ultimately from a circuit court decision based on case information and (at least in some part) on recommendations of the supervising probation officer. For technical violations and absconder warrants, officers in consultation with their supervisors have some discretion in determining whether the evidence of probation failure is serious enough to bring to the attention of the court. But the courts have the final decision on how a case of probation violation will be disposed.

In this study, risk assessment is based on the estimated probabilities that a probationer will fail by one of the three defined modes during any one of a non-overlapping sequence of six-month intervals. The probabilities are based on a statistical examination of patterns of failure among all probationers admitted between January 1, 1991, and December 31, 1994[4]. Case outcome information (release, failure, etc.) was complete through the end of May, 1995.

As described in more detail in Section 2, statistical models were estimated to identify these probabilities for each of the three failure modes and for each of the intervals beginning with months 4, 10, 16, 22, 28, and 34 after admission. The models assume, of course, that a subject probationer is still in active supervision status at the beginning of the interval. Probabilities were also calculated for two other periods: the first three months of supervision and the nine-month interval from month 40 through the end of 4 years of supervision. In implementing the risk assessment, probabilities generated by the model for the initial 3 month period will not be used for classification because of the the inevitable lag between intake and complete case record availability. Rather, once the offender data are available in the management information system, the initial classification will be based on the probabilities derived here for failure in the interval from the beginning of month 4 through the end of month 9.

In practice what this classification system implies is that the assessed failure probability of a probationer who remains in active status will change at specific points over the course of his sentence. These probability changes will presumably be large enough to

[4]In total there were about 184,000 cases available for the analyses carried out in this study.

result in a risk level reclassification of some fraction of the supervised population--generally, although not always, a reclassification to a lower risk level.

For each case and for each of the time periods, the probabilities estimated for the three different failure modes can be combined in a number of different ways to give an overall risk score. For example, suppose a probationer had probabilities of 0.15, 0.08, and 0.03 of being revoked for a new arrest, revoked for a technical violation, or absconding, respectively, during a particular six-month period. If policy dictates that each of these outcomes is to be regarded as equally serious, a risk score that gives equal weight to each outcome probability can be calculated by simply adding the three probabilities to yield a probability of 0.26 for overall risk of failure within that interval. However, a policy maker might believe that revocation for a new arrest (reflecting allegations of new criminal activity) is more serious than either a technical violation or absconding. Policy might then dictate that the risk of revocation for a new arrest be weighted more heavily[5]. In this report, the three probabilities are given equal weight and the overall risk score is calculated as the sum of the probabilities of failure during the interval of interest.

There are many individual characteristics that might be linked theoretically to success or failure on probation. They include, for example, age and gender, criminal record, drug or alcohol addiction, economic status, employment history and ties to family or the community. However, much of this personal information is not routinely collected and automated and, thus, would be expensive to collect and difficult to measure with accuracy and reliability. In developing a risk-scoring system to be applied to all active cases, it was assumed that the assessment should be based on data that were, at least for most cases, already recorded in the Department of Corrections computerized information system. The following very limited list of data elements was chosen for this study:

---

[5]Only if the failure modes have equal weights is their sum interpretable as an overall <u>probability</u> of failure. Otherwise the weighted sum of failure probabilities establishes a risk scale determined both by the individual failure mode probabilities and the dictates of policy regarding the relative seriousness of the different possible outcomes. The distinction in <u>interpretation</u>, however, is irrelevant for purposes of classification.

- **Personal characteristics**
  - Gender
  - Age at admission
- **Prior criminal history**
  - Number of prior Florida prison commitments
  - Number of previous admissions to community supervision
- **Current offense**
  - Classification of most serious conviction charge as personal violence, property crime, drug offense, or other
  - Number of counts
  - Length of probation sentence
  - Split sentence
- **County of supervision**
  - Judicial Circuit
  - DOC Region

Although most of these variables are obvious candidates for inclusion in a study examining patterns of probation failure, a few words of explanation may be in order for introducing jurisdictional variables derived from the county of supervision. In an early, exploratory study, we estimated models without including jurisdictional indicator variables. The results from these models were satisfactory when the entire state was considered as a single jurisdiction (i.e., observed numbers of failures compared favorably with expected). But when we attempted to predict outcomes for cases restricted to one of the five Department of Corrections administrative regions or to one of the 20 state judicial circuits, the results simply did not fit the observed data. Many reasons might be suggested to explain why ostensibly similar groups of cases from different jurisdictions should on average have different outcomes. But in essence what is demonstrated in these findings is that the courts' discretionary decisions to revoke probation or to issue an absconder warrant vary among the different circuits and, presumably, so do the decisions on the part of the probation offices to recommend revocation. Additionally, offenders and their criminal behavior may vary systematically across these jurisdictions in ways not captured by the limited set of variables included in the models. Indeed, given differences in local "court culture," in caseloads and in the nature of cases being supervised, there seems to be no *a priori* reason to expect uniformity across the state in offender behaviors or in policy toward these behaviors leading to an officially defined "failure."[6]

1.3. RISK-BASED CLASSIFICATION

Application of the risk assessment methods described in this report results in the assignment of a failure probability, a number between 0 and 1, to each probationer currently under active supervision. This risk score represents the probability that the probationer will be officially declared a failure by any of the three modes during the current assessment interval (e.g., months 4 through 9 of supervision) and is generated by summing the probabilities of each negative outcome.

Once classified, any group of probationers can be ranked by risk score. This ranking can then be used to establish a triage in which, for example, a subject is classified as high, medium or low risk. To go from this continuous risk scale to a three level classification requires the introduction of two "cut points." All subjects with failure probabilities greater than the upper cut point are classified as high risk; those with probabilities less than the lower cut point are the low risk cases. Obviously, those with probabilities between the two cut point levels are the medium risk class.

The placement of these cut points is a policy decision, albeit one that is constrained by resources. Concerns over public safety would suggest relatively low values for both

---

[6]We emphasize again that "failure" and its associated risk are defined in this study by court orders and thus capture both the behavior of probationers and processes and decisions of criminal justice system officials. In principle one might base the risk assessment on a different definition of failure (and time to failure)--for example on the fact of an arrest or on a probation officer's report of a technical violation or absconding. The statistical methods used in this study would still be applicable.

cutpoints: few subjects classified as low risk and low threshholds for defining the higher risk classes. This would tend to place individuals in relatively restrictive classifications and would reflect a supervision policy that is reluctant to take chances. Conservation of supervision resources would dictate a diametrically opposite policy--relatively high values for both cut points, which would tend to identify relatively few high risk cases with the maximum acceptable number classified low risk. Thus, the decision as to where to place these classification cut points necessarily involves a trade-off between the demands for greater supervisory control of offenders and the real constraints that are dictated by the workloads implied.

Figures 1 and 2 illustrate how the population classification might occur under two hypothetical cut point placements. In each figure we are concerned with a cohort of individuals whose numbers diminish over time due to failure or release. Individuals still in active supervision status at the beginning of each six month interval are scored according to the model for that interval, ranked according to this score, and classified as noted. The figures show the percentage of the population assigned to each of the three classifications during each six-month interval. The figures also show the associated failure rates.

In Figure 1, all subjects whose overall failure probability was 0.10 or less were classed as low risk; those with failure probabilities greater than 0.25 were considered high risk cases. Given the range of failure probability values estimated from the data used for this study, this would represent a fairly conservative classification policy. Figure 1 reveals the following:

1. The fraction of the admissions cohort that is classified as low risk increases over time from about 12% to about 55% of those probationers still under active supervision. Conversely, the fraction of the cohort classed as high risk decreases over time. Initially, about 18% of the population is classed as high risk but this decreases to

about 3% of the long term survivors. The fraction classified as medium risk remains quite large, decreasing from about 70% to about 40% of active cases.[7]

2. The observed, six month failure rates under this hypothetical classification remain

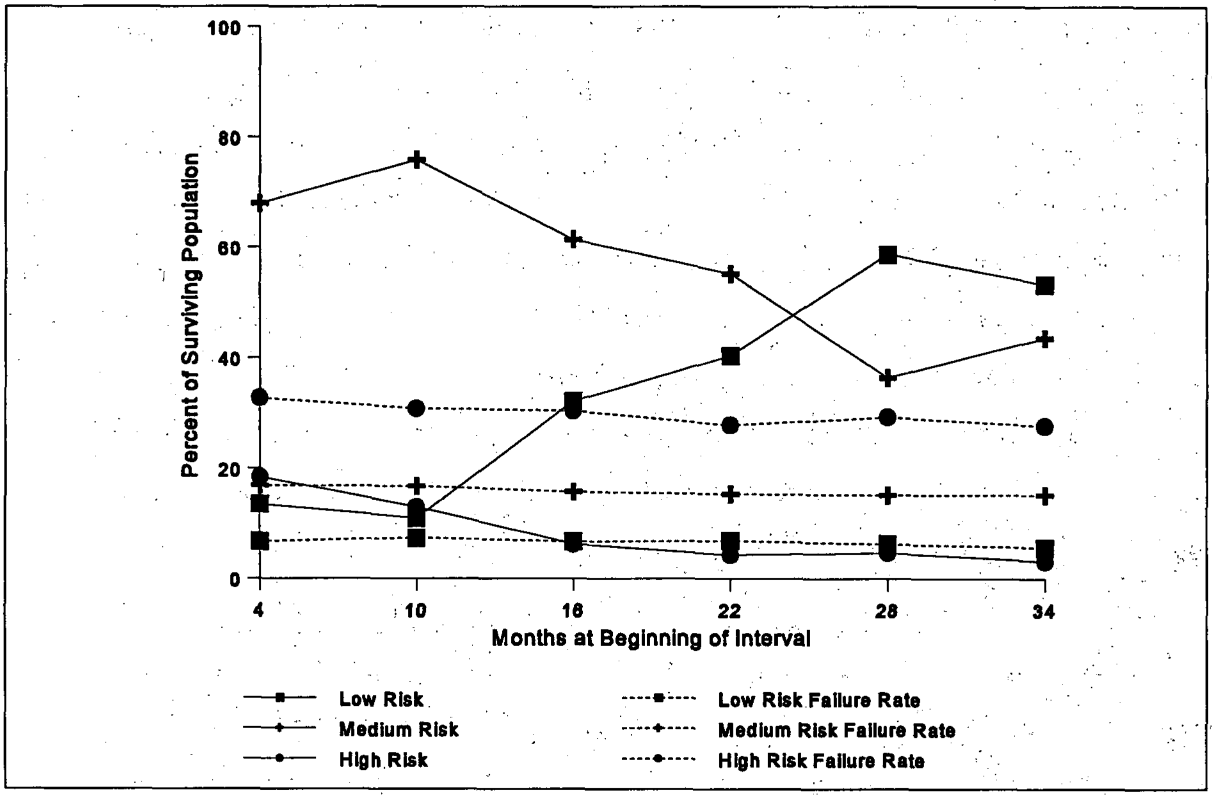


Figure 1. Example of a "Conservative" Classification Scheme. Low Risk is defined as a failure probability of < 0.10 and high risk is defined as a failure probability > 0.25.

fairly constant for each of the three risk classifications. About 6 or 7% of active, low risk cases, about 16% of medium risk cases, and 30% of high risk cases failed in each time interval.

---

[7]Note that these are percents of populations of very different sizes. Among active cases in late summer 1995, about 31,000 had been under supervision for less than nine months; only about 3,300 had successfully completed between 34 and 39 months of their sentence. See Section III for further discussion.

Thus, the models appear to be fairly successful in defining a risk hierarchy in which those classified as low risk are considerably less likely to fail than those classified as medium risk, who are again less likely to fail than the high risk cases.

Figure 2 illustrates a hypothetical classification policy that assumes more severely constrained resources. Specifically, we wish to assign relatively more cases to minimum supervision and relatively fewer cases to maximum supervision. Here all active cases with failure probabilities below 0.20 are classed as low risk; probationers with probabilities greater than 0.35 make up the high risk class. Initially, about 66% of active cases are classed as low risk, increasing over time to over 90% of the surviving population. The medium risk class is much smaller than under the divisions of Figure 1, decreasing monotonically in time from about 28% to about 8% of the long term survivors. Finally, only relatively extreme cases are classed as high risk: initially about 6% of an intake population but decreasing to less than 1% of long term survivors. Under this less restrictive classification scheme, as one would expect, we observe higher fail rates by class. The observed six-month failure rates decrease over time from about 14% to about 9.5% of active low risk cases. The rate among medium risk cases remains fairly constant at 24 or 25% while failure among the relatively small class of high risk cases decreases from an initial 42.5% to about 35%.

It should be noted that in each risk class the failure rates under the less restrictive classification policy of Figure 2 are considerably higher than under the more conservative policy illustrated in Figure 1. This result follows from the addition of cases with higher failure probabilities into the lower risk categories in the less restrictive or "constrained resources" classification. Specifically, for example, in Figure 2, the low risk class is defined to include not only those with a failure probability of less than 0.10 but also those whose failure probability is greater than 0.10 but less than 0.20. Thus, failures for the "low risk" group would be expected to be higher.

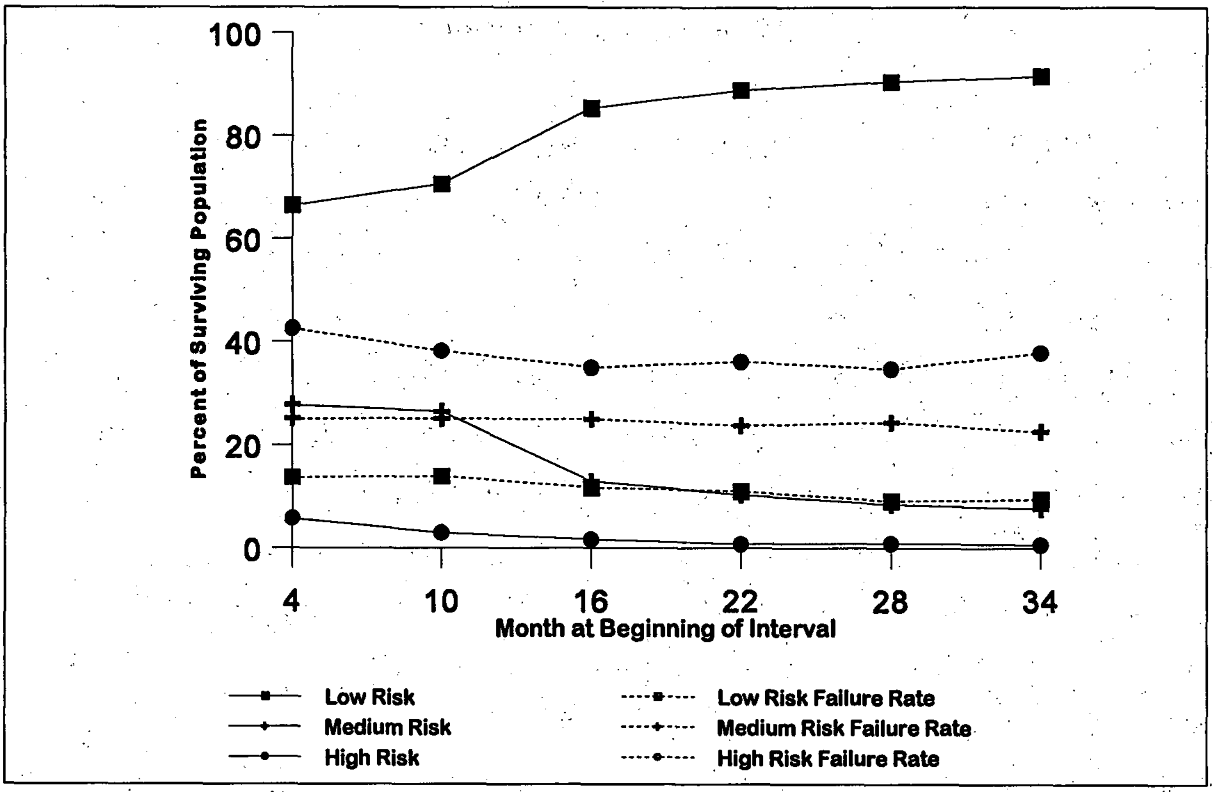


Figure 2. Example of "Constrained Resources" Classification Scheme. Low Risk is Defined as Failure Probability < 0.20 and High Risk is Defined as Failure Probability > 0.35.

The two examples shown in Figures 1 and 2 illustrate classification policies assumed to be uniform across the entire state. It would add only minor complications to a classification system to allow the classification cut points to vary among regional offices, among judicial circuits or even among local probation offices. Such a jurisdiction-specific classification policy would make sense if, for instance, there were substantial variation in case loads over the state.

2.0 MODEL DEVELOPMENT

In developing and applying a risk classification system to serve as a basis for identifying levels of supervision and allocating probation officer resources, it is important to specify the length of the time interval over which the classification is valid. While the probability that an individual will eventually fail in the course of his sentence may be of interest in some applications, the probation officer is presumably more interested in the practical question of the current risk of near term failure for each probationer on his caseload and how these risks are distributed among the subjects for whom he is responsible. Of course, these two probabilities--failure ever during a sentence and failure during some specified, near term time interval--are related. However, all other things being equal, risk measured by likelihood of ever failing will increase monotonically with sentence length simply because of the greater length of time over which the subject with the longer sentence is at risk. Conceivably, then, a classification based on total supervision time could result in all probationers with "long" sentences being defined to be high risk while subjects with "short" sentences are classed as low. From the perspective of forecasting failure in the relatively near future, such a risk ranking might actually turn out to be perverse[8].

Figure 3 shows the distribution of sentence lengths rounded to 0.1 years for sentences less than or equal to five years. The sharp spikes in this distribution correspond to probation sentences commonly handed down: six months, one year, eighteen months and so on.

Figures 4a through Figure 7a show the distributions of times to revocation for a new arrest, revocation for a technical violation, issuance of an absconder warrant and release. Figures 4b through 7b show the corresponding distributions of the ratios of time to case disposition divided by length of the probation sentence imposed at intake. The distributions of times to failure as a fraction of sentence (Figures 4b through 6b) indicate that case

[8]Indeed, as will be seen later in this report, six month failure probabilities generally decrease with increasing sentence length.

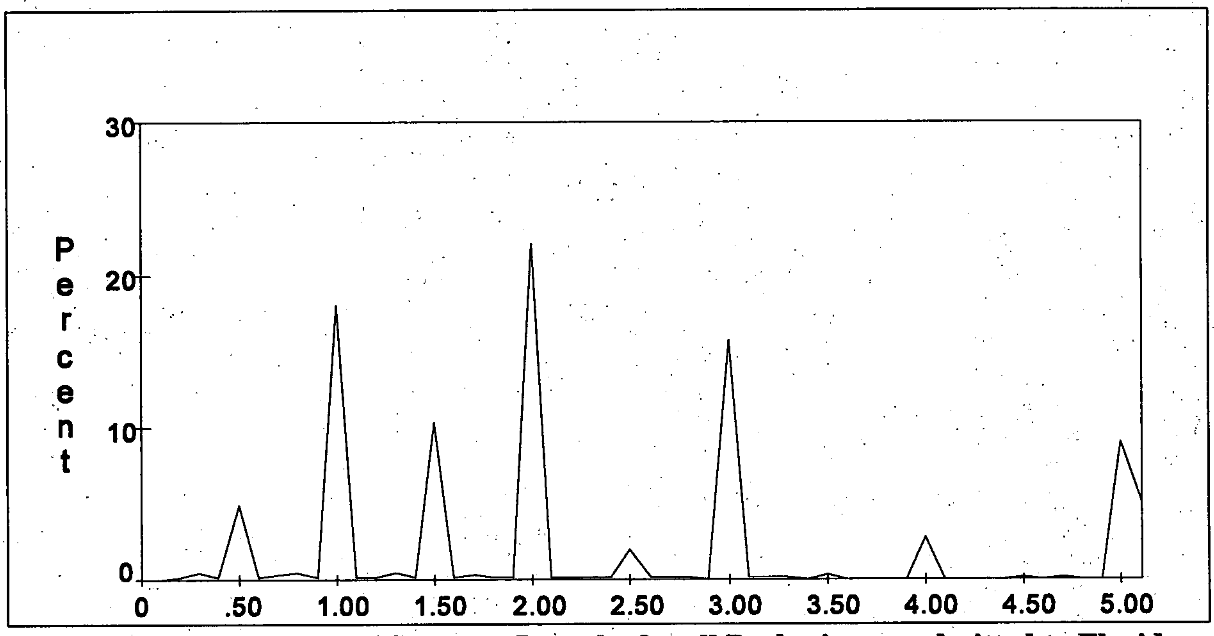


Figure 3. Distribution of Sentence Lengths for all Probationers admitted to Florida Probation between January 1, 1991, and December 31, 1994.

disposition actions can occur throughout a sentence although many apparently are delayed until the sentence is about to end.

In cases of technical revocations or absconder warrants, the probation office or the court may want to give a subject as much time as possible to satisfy conditions of his sentence (perhaps payment of a fine or restitution) or to return voluntarily from absconder to active supervision status. Revocations for a new arrest that occur at the end of the probation sentence are most likely an indication that the times to case disposition may be influenced by crowded court dockets.

As can be seen in Figure 7b, release most often occurs at the completion of an assigned sentence. However, both early release and "late" release are possible. If progress under supervision is considered satisfactory, the supervising probation office may refer a case to the court with a recommendation of early release. Conversely, probationers may be continued under supervision beyond their initially established release date if, for example, they were in absconder status for a period of time and the court decides on a corresponding delay in the release date.

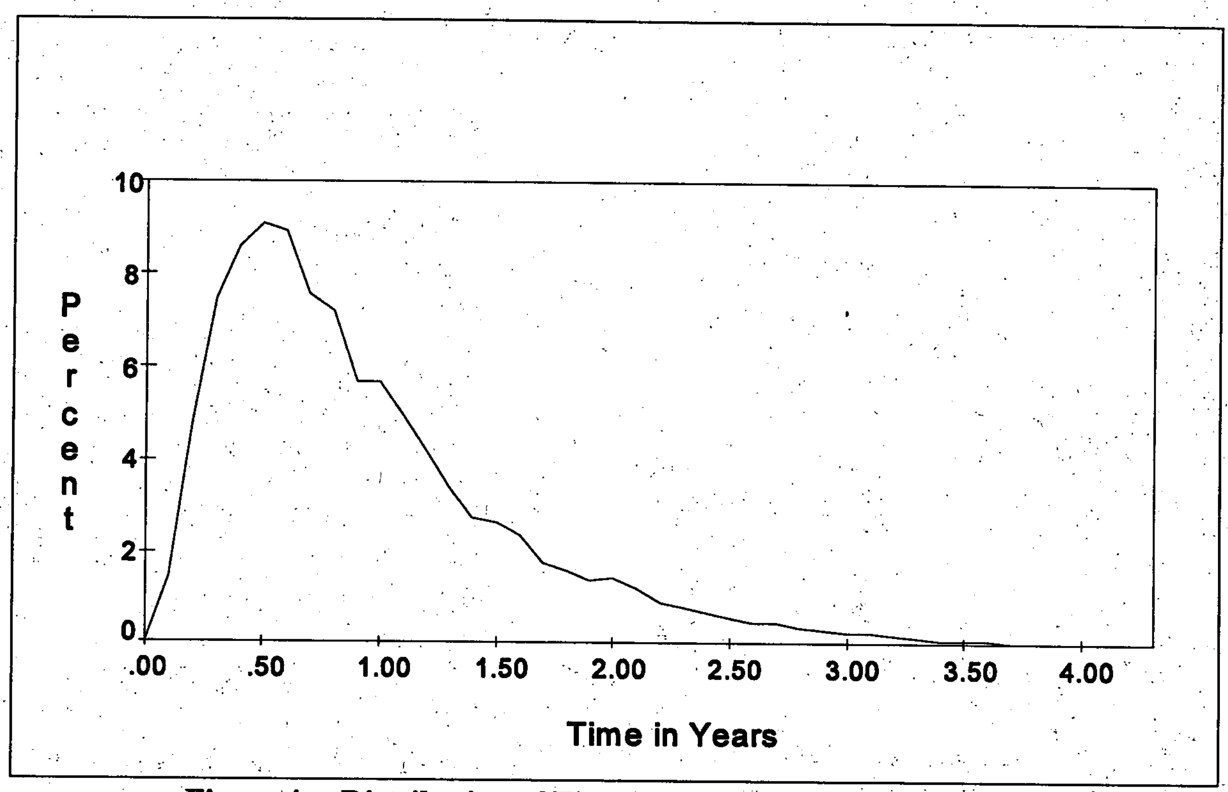


Figure 4a. Distribution of Time to Revocation for New Arrest.

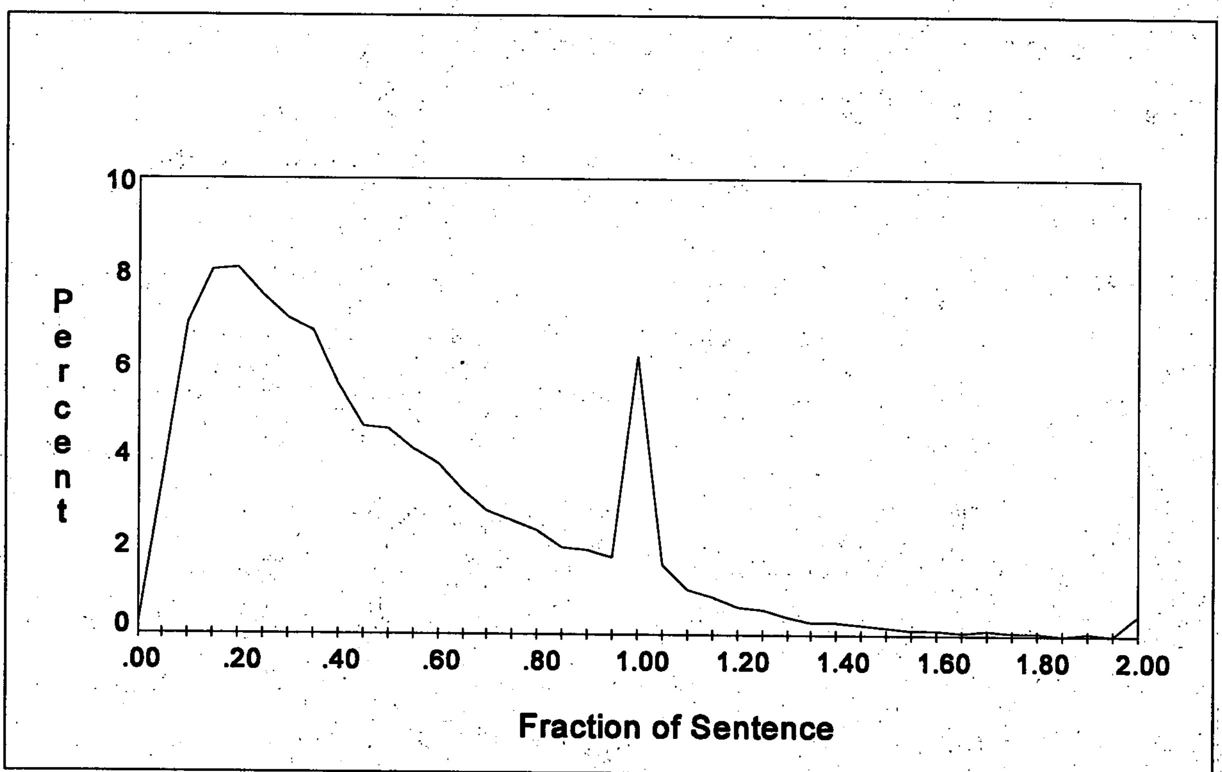


Figure 4b. Distribution of Time to Revocation for New Arrest as a Fraction of Sentence.

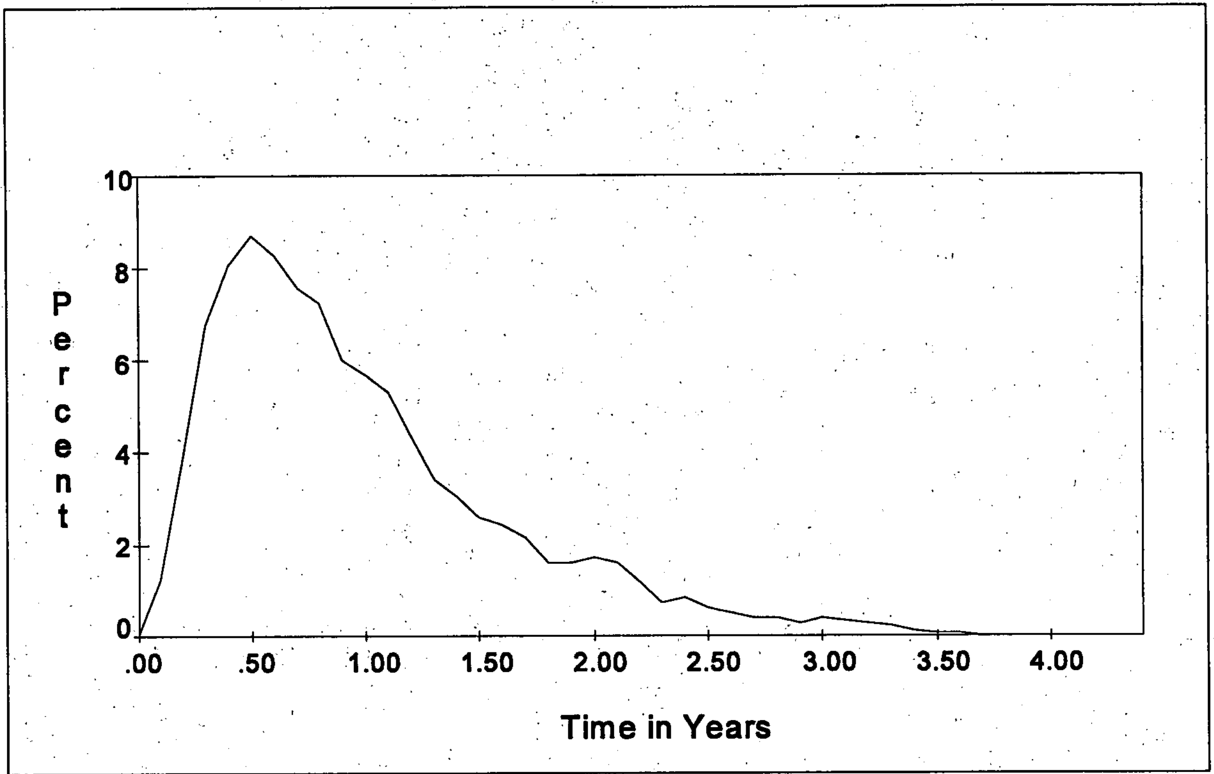


Figure 5a. Distribution of Times to Revocation for a Technical Violation.

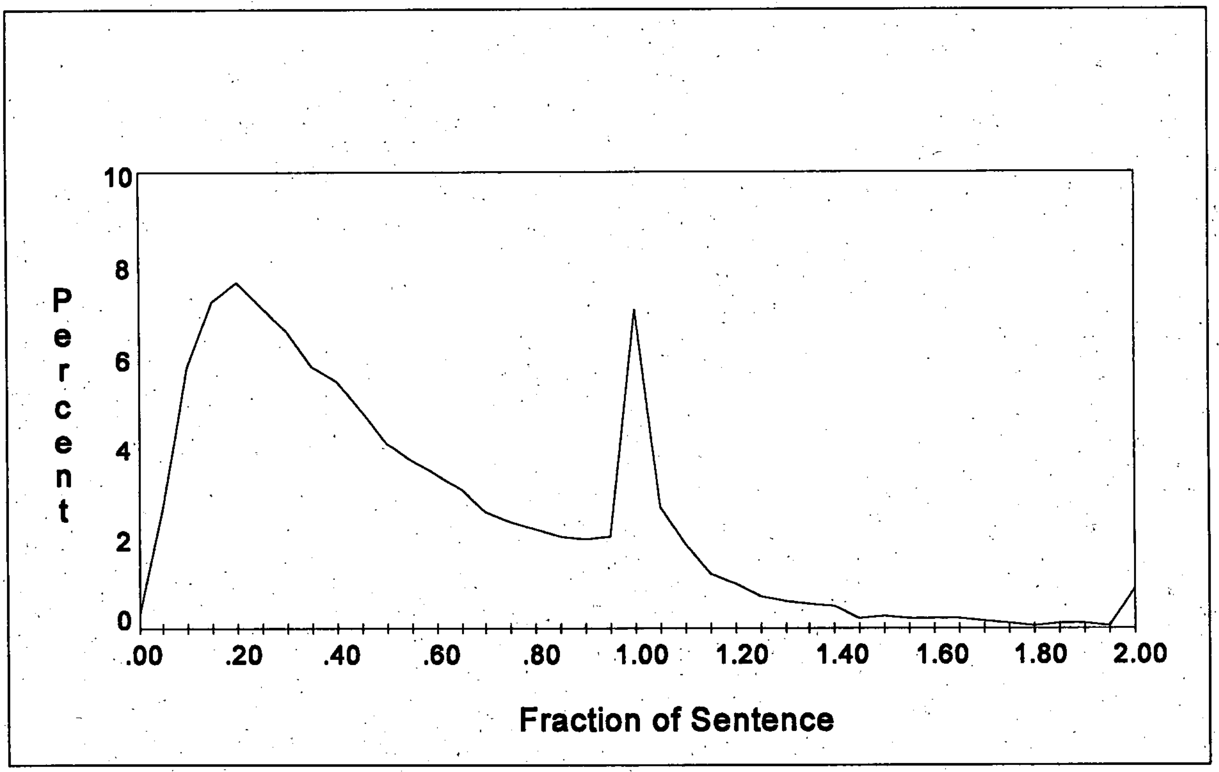


Figure 5b. Distribution of Time to Revocation for Technical Violation as a Fraction of Sentence

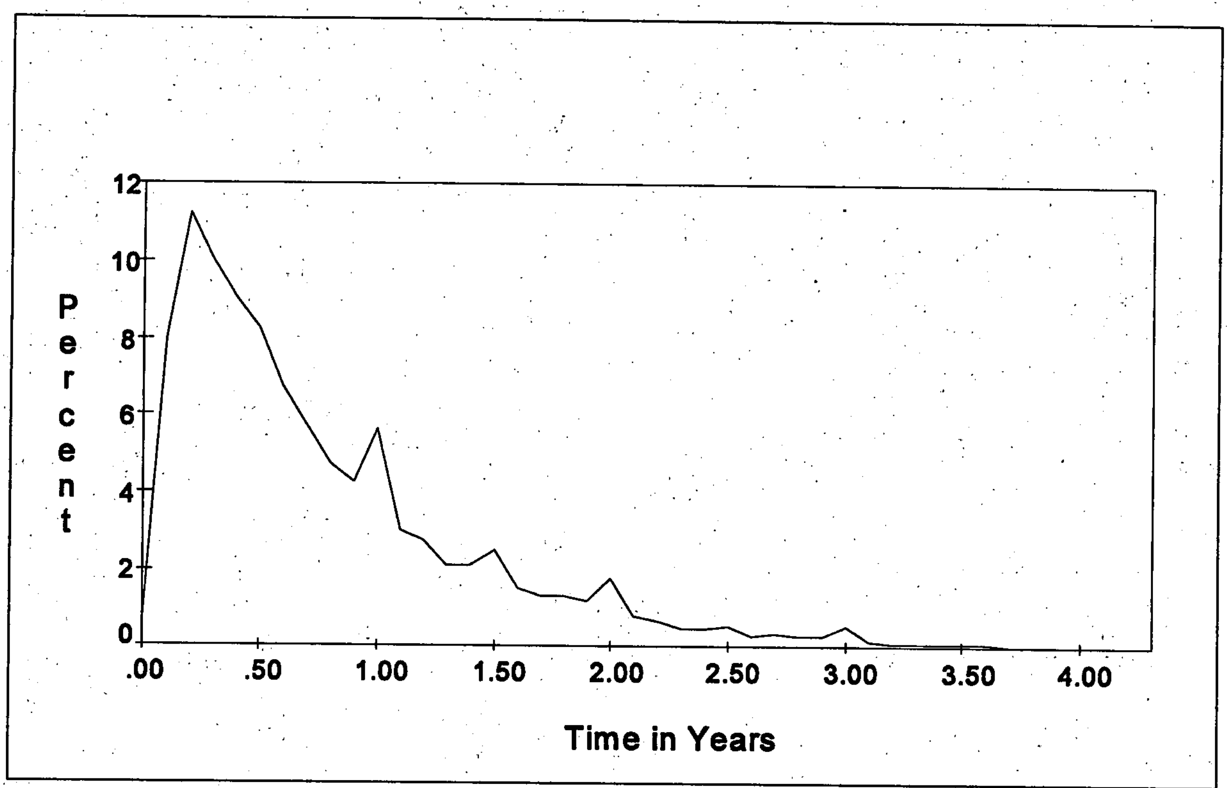


Figure 6a. Distribution of Time to Absconder Warrant.

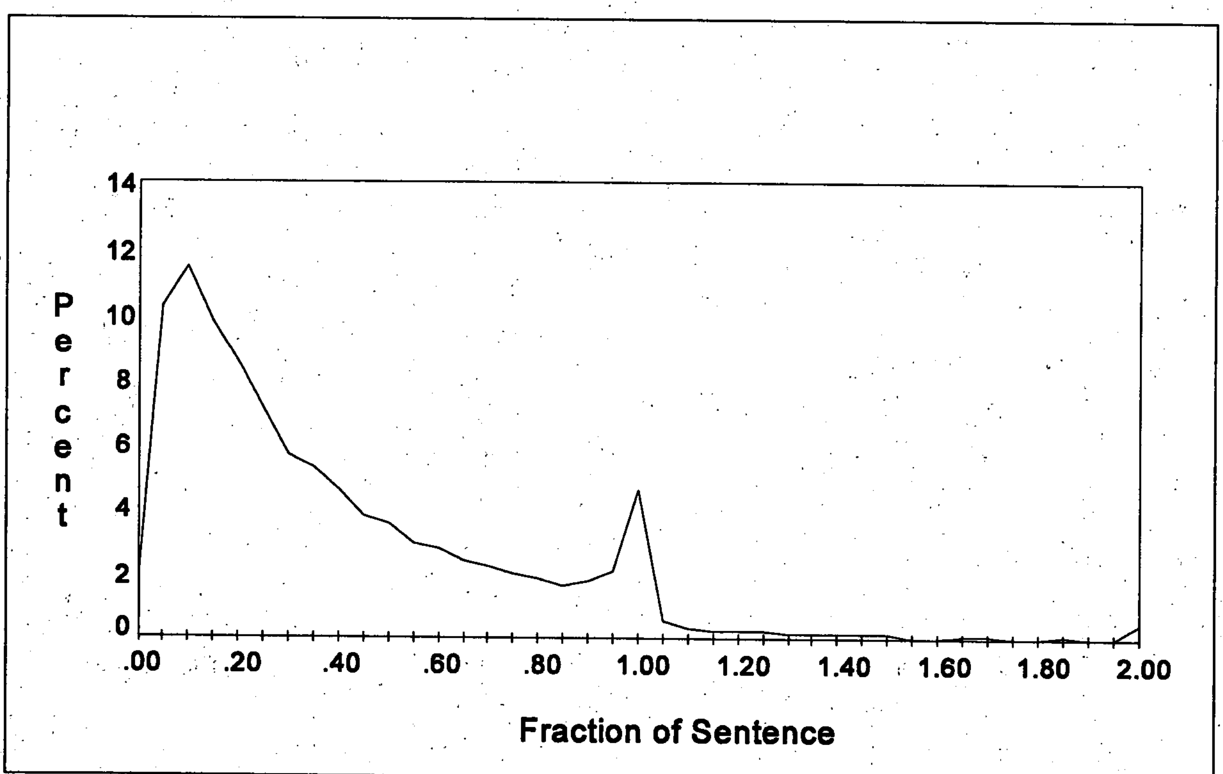


Figure 6b. Distribution of Time to Absconder Warrant as Fraction of Sentence.

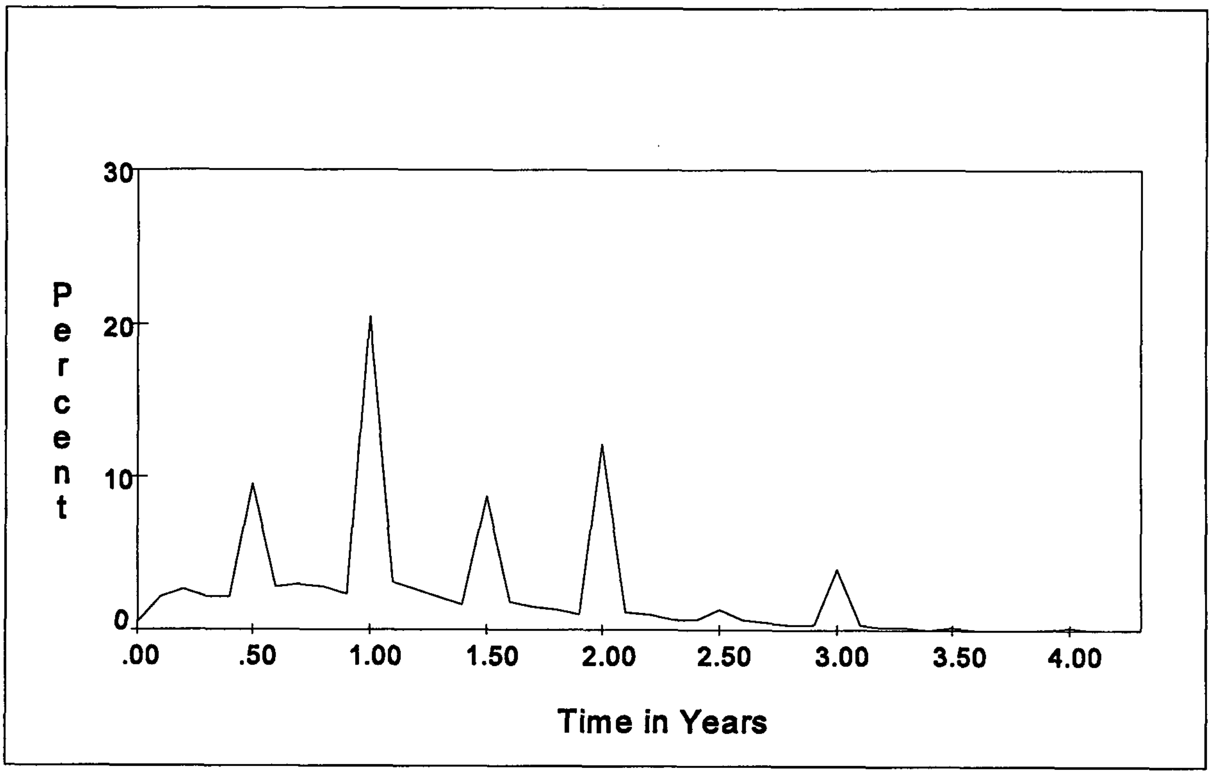


Figure 7a. Distribution of Time to Release.

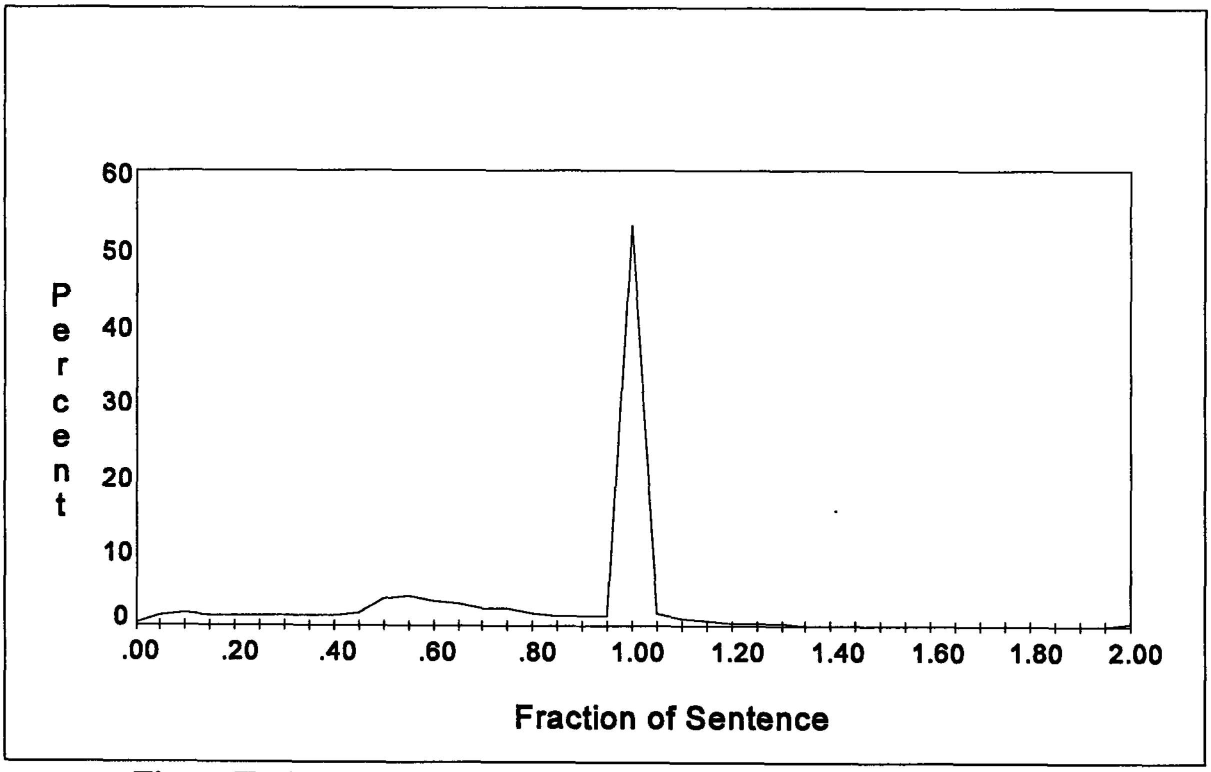


Figure 7b. Distribution of Time to Release as a Fraction of Sentence.

2.1 STRUCTURE OF THE PROBABILITY MODELS

In this section, we develop the general mathematical structure of the probability models chosen as the basis for a risk classification. In subsequent sections, the independent variables used to characterize each case are defined and the principal features of the fitted models are explained.

For this classification problem, we are interested in both whether a specific event occurs and when that event occurs. One type of model that is appropriate for this type of problem is the *hazard* or *survival* model. The fundamental concept underlying this statistical approach is the hazard function--the conditional probability of observing the outcome of interest in a short time interval, given that the subject is still under active supervision at the beginning of the time interval. If there is more than one potential outcome (as there is in this study), a *competing hazards* model is used. The *survival function* is the probability that a subject "survives" or, in the case of our problem, is still actively on probation at time t. In this formulation then, successful completion of a probation sentence (release) is necessarily one of the competing, modeled outcomes along with the failure modes. Because of their flexibility in allowing one to make predictions for any specified time interval, hazard models offer a particularly powerful approach to the probation risk classification problem. Unfortunately, our efforts to fit a competing hazards model to the Florida probation data were unsuccessful. Specifically, it appears that the mixture of cases actually terminated on a pre-specified date with cases in which time to termination might be considered a random variable defeated our attempts at a hazard model formulation[9].

---

[9]Hazard models have been used successfully in recidivism studies in which some measure of individual failure risk is assumed to continue indefinitely in time. With probation failure, of course, the risk drops discontinuously to 0 at the time of release. The probation problem would be tractable within a hazard model formulation if, for example, failure were defined stochastically as occurring at the time of arrest or of a probation officer's noting some other violation rather than at the time of a court's decision and if release always occurred at the termination of the sentence imposed.

As an alternative to the hazard model approach, which divides the time line after admission into infinitely small segments, we decided to model probation outcomes over a non-overlapping sequence of finite-time intervals. Models were developed for months 0 through 3, 4 through 9, 10 through 15, 16 through 21, 22 through 27, 28 through 34, and 35 through 44. Outcome probabilities, conditioned on a subject's being in active supervision status at the beginning of an interval, were determined using a *multinomial logit model.*[10]

With the multinomial logit model, we are interested in identifying the relative odds of various outcomes. Suppose that in any specified time interval there are K possible outcomes--including the reference outcome "still active at end of interval." Then,

$$\ln\frac{p(k|X_i)}{p(K|X_i)} = \sum_{j=1}^{J} X_{ij}\, b_j^{(k)} \tag{1}$$

where X_{ij} is subject *i's* value for the j^{th} variable and $b_j^{(k)}$ is the model coefficient for variable *j* corresponding to the k^{th} outcome. In these expressions, there are J independent variables (including the intercept). Equation (1) expresses the basic distributional assumption of this form of multinomial logit model: The log of the odds of outcome *k* relative to the reference outcome *K* is a linear function of the explanatory variables.

Initially, five distinct outcomes were modeled: Revocation for rearrest, revocation for technical violation, abscond[11], release and the reference outcome, still active at end of the time interval. Again, as with the competing hazards model, goodness-of-fit tests suggested that the results for some of these models still did not give an adequate fit to the outcome data. However, satisfactory results were obtained when we combined "released" and "still active" into a single outcome. Consequently, in the results reported here, the models are for failure

---

[10]For ready reference a brief discussion of logistic regression models is contained in Appendix A. Appendix B provides some information on model estimation software.

[11]Cases in which revocation of probation followed return from absconder status were defined as absconders and the failure date was taken as the date the absconder warrant was signed.

vs. non-failure in each interval, with failure defined by three distinct outcome modes and non-failure as "released or still active."

2.2 DATA AND MODELS

The data base for the model estimation included all probation admissions over the four-year period from January 1, 1991, through December 31, 1994. Case outcome information was complete through the end of May 1995. A time-at-risk based sequence of eight separate models was estimated, with each model using data from the set of most recent cases that satisfied two conditions:

1) The case was still active at the beginning of the interval; and
2) The outcome at the end of the interval (failure vs. non-failure) was known.[12]

Model 1 covers the first three months after intake, with a construction sample of about 45,000 observations consisting of all intakes during 1994. Model 2 covers the period from the beginning of the fourth month after admission through the end of the ninth month. Cases selected were all those admitted after July 1, 1993, and active at least 91 days after intake (N = 50,000). Model 3 covers the next six-month period (months 10 through 15), drawing its data from all cases admitted since January 1, 1993, and under active supervision for at least 273 days (N = 38,000). Models 4 through 7 each cover similar six month intervals, terminating at the end of month 39. The numbers of cases on which these model are based decreased from about 38,000 for model 4 to about 10,000 for model 7. The final model covers the nine-month period from the beginning of month 40 through the end of month 48 (N = 2400).

---

[12]The second condition was necessary because, unlike hazard models, logit models cannot handle observations for which the outcome during a time interval is not known (i.e., censored observations).

Observed six-month failure rates generally decline over successive time intervals. Rearrest rates among active subjects decreased from 0.061 during the period from months 4 through 9 to 0.027 in months 34 through 39. Somewhat smaller changes were also observed in technical revocations (0.056 to 0.035) and in absconding rates (0.067 to 0.044).

All variables characterizing a particular case were contained in the Department of Corrections' management information system although, as used in the models, some transformations of the raw data were made. The following list contains the data elements used for these analyses.

- **Personal variables and criminal histories**
  - Sex: (0 = male; 1 = female)
  - Ln (age in years at admission - 17) [LAGEADM]
  - Prior Florida prison terms [PRPRSN]: Integer
  - Prior Florida community supervision [ADMITS]: Integer
- **Variables characterizing most serious offense and sentence for current conviction**
  - Violent offense: (0 = no; 1 = yes)
  - Drug offense: (0 = no; 1 = yes)
  - (Property offense = reference category)
  - Other offense: (0 = no; 1 = yes)
  - Split Sentence: (0 = no; 1 = yes)
  - Ln (Number of counts truncated at 25) [LCOUNTS2]
  - Ln (Sentence length in years truncated at 25) [LYRSUP2]
  - Probation sentence at conviction scheduled for completion in this or a previous model interval [CATSUP]: (0 = no; 1 = yes)
- **Jurisdiction indicators**
  - Four indicator variables to specify to which of the five DOC regions the case is assigned for supervision

- Fifteen indicator variables to define the circuit court of supervision (total of 20 circuits)

Truncations and log transformations were introduced to reduce the mathematical effects of extreme values in certain variables. For the age-at-admission variable (LAGEADM), there were no juvenile cases (age less than 18) in the data base used for model estimation although a few such cases are remanded by state courts to probation supervision. The modal age was 19; but 0.3% of the study population was over age 70 at the time of admission. By subtracting 17 from the age in years, we scaled this variable so that its log (LAGEADM) would take the value 0 for the youngest members of the dataset (i.e., those 18 years of age).

For the number of counts (LCOUNTS2), we observed that among the 184,000 cases in the data base, 65% were convicted on a single count. The distribution of counts drops off rapidly with 99% of cases charged with 10 counts or fewer. However, the very long tail of this distribution includes cases with hundreds of counts--the maximum being 992.

For sentence length (log sentence length in years = LYRSUP2), again the distribution has a long, thin tail. Ninety-nine percent of cases were sentenced to less than 13 years of probation, but 126 out of 184,000 cases received sentences of 90 years or more.

Florida counties are aggregated into twenty judicial circuits and five Department of Corrections administrative regions. In these analyses, the region in which a case is supervised was characterized by four indicator variables with Region 5 chosen as the reference region. Within each of the five regions, one circuit was designated as the reference circuit with the remaining circuits again characterized by indicator variables. Thus, for example, if the county of supervision is located in Circuit 2, which is in Region 1, the case would be coded as CIRCT2 = 1, REGION1 = 1 with 0's for all other circuit and region variables. If the county of supervision were in Circuit 1, the reference circuit for Region 1, it would be coded as REGION1 = 1 with all other region and circuit variables coded 0.[13]

The probation sentence is described by two variables in these models. The log of years of supervision is, of course, continuous and monotonically increasing. A second variable, CATSUP was introduced as a step function that takes on the value 1 in the period in which the sentence was scheduled to end at the time of admission and in all subsequent periods. In intervals prior to this the variable has the value 0. This variable was introduced to reflect in some measure the discontinuities found in the distribution of outcomes since releases and, to a lesser extent, revocations and absconder warrants tend to occur at the end of the imposed sentence. This variable was not included in the model for the initial 3 months of supervision.

As noted earlier, we used the most recent cases available that provided sufficient followup time to estimate the models. Thus, each model was based on partially overlapping but different sets of observations. Other than the variables describing the sentence length, however, the population means do not change very much between the recently admitted cases used for modeling risk in the first three months and those longer-term cases used to estimate subsequent models. The female population remains roughly constant at about 20%. Between the first and eighth datasets, the fraction of the surviving population serving sentences for a violent crime increases from about 21% to about 27%, while drug offense probationers decrease from 27% to 21% of the population. The representation of split sentences increases from 5% to 11%.

Interestingly, the population means of the two criminal history variables both decrease as the population evolves to include more individuals with lengthy sentences. Specifically, the average number of prior prison terms drops from 0.21 to 0.15 and the average number of prior admissions to community supervision drops from 0.57 to 0.50. Thus, the chances of long-term survival in active supervision would appear to decrease with

[13]In the model for the last time interval, circuits 16 and 19 were combined into a single variable because of the small number of long term cases found in the data for these circuits.

increasing severity of prior criminal history as measured by these two variables. This could be explained by higher failure rates among offenders with prior Department of Corrections commitments.

Although this study was not designed to investigate variability in probation sentences, it is of some interest to examine the extent to which sentence length is "explained" by the other independent variables used in the failure models. Table 1 gives the results of a linear regression of sentence length on the remaining variables. Because of the large number of cases available for this analysis (N = 183,821), almost all of the coefficients are determined with great precision and, hence, the regression can be considered to give good estimates of the expected value of the length of probation sentences, given the values of the other independent variables. On average, sentences increase with increasing prior prison commitments (PRPRSN) but decrease with number of prior admissions to community supervision (ADMITS). One might speculate that in the sentencing decision a previous commitment to prison is correlated with or at least regarded as an indicator of serious criminality whereas previous community supervision commitments reflect repeated but relatively minor offending. In any case it will be noted that the variance in lengths of sentences is quite large and that a linear relation between sentence length and these twenty-eight regressor variables explains only about 6 percent of this variance.

Table 1. Regression of Sentence Length on Other Independent Variables

Variable	B	SE (B)	Beta	t	p > |t|
VIOLENT	0.581679	0.022203	0.066338	26.199	0.0000
CIRCT5	-0.292078	0.049279	-0.016494	-5.927	0.0000
PRPRSN	0.311014	0.014550	0.057056	21.376	0.0000
CIRCT16	0.217272	0.078957	0.006456	2.752	0.0059
CIRCT8	0.317750	0.067930	0.012463	4.678	0.0000
CIRCT14	0.948856	0.063576	0.040888	14.925	0.0000
CIRCT19	0.402465	0.050549	0.019867	7.962	0.0000
CIRCT3	1.495055	0.075710	0.050874	19.747	0.0000
CIRCT20	0.596670	0.051450	0.030008	11.597	0.0000
LCOUNTS	0.736777	0.014964	0.115640	49.236	0.0000
CIRCT2	0.196772	0.056295	0.010255	3.495	0.0005
CIRCT12	-0.206217	0.048994	-0.011059	-4.209	0.0000
SEX	-0.119000	0.020450	-0.013306	-5.819	0.0000
CIRCT10	0.966015	0.047190	0.054606	20.471	0.0000
CIRCT7	1.420938	0.054730	0.077700	25.963	0.0000
LAGEADM	0.194959	0.008940	0.050448	21.807	0.0000
CIRCT15	0.433240	0.044090	0.025403	9.826	0.0000
CIRCT18	0.110181	0.047368	0.006587	2.326	0.0200
OTHER	-0.375844	0.027711	-0.032961	-13.563	0.0000
CIRCT13	0.606881	0.037242	0.049911	16.296	0.0000
SPLIT	1.541400	0.034603	0.106344	44.546	0.0000
CIRCT11	-0.616624	0.034092	-0.053066	-18.087	0.0000
DRUG	-0.202230	0.019872	-0.025843	-10.176	0.0000
ADMITS	-0.150033	0.008004	-0.050892	-18.744	0.0000
REGION3	0.620179	0.039783	0.064449	15.589	0.0000

Variable	B	SE (B)	Beta	t	p > |t|
REGION1	-0.148115	0.045183	-0.012929	-3.278	0.0010
REGION2	-0.393329	0.045425	-0.036198	-8.659	0.0000
REGION4	-0.110980	0.034967	-0.014552	-3.174	0.0010
(Constant)	1.725707	0.034903	-	49.443	0.0000

Note: R-square = 0.06382; F = 447.50000, p-value (F) = 0.0000

2.3 LOGIT MODEL RESULTS

In this section, we discuss the results of a "typical" one of the eight multinomial models estimated for this project. (Results from the other seven models are provided in Appendix C.) For all of the models,

- Outcome 1 is revocation for a new arrest during the time interval;
- Outcome 2 is revocation for a technical violation;
- outcome 3 is abscond; and
- Outcome 6 refers to the "no failure" outcome--release or still active at the end of the interval.

Table 2 reproduces the results for Model 4, the model covering the interval from the beginning of month 16 through the end of month 21. Positive values of coefficients associate increasing variable values with an increase in the odds of a particular failure outcome relative to "no failure." Thus, other things being equal, the failure odds of female probationers still under active supervision after 15 months are lower than those of males during the interval considered here. Similarly, failure odds decrease monotonically with increasing age at admission (LAGEADM). Split sentences, longer criminal histories and a greater number of counts all tend to increase failure odds. Offenders convicted on violence or drug charges are somewhat more likely to be revoked than property offenders but marginally less likely to abscond; offenders whose most serious conviction charge is in the "other" category have lower failure odds than property offenders. Finally, probationers serving longer sentences (LYRSUP2) have lower failure odds but this result is also dependent on whether their sentence is scheduled to end during this time interval or in a previous one--that is, on the value of the variable CATSUP.

TABLE 2. LOGIT MODEL RESULTS: MODEL 4 (MONTHS 16 THROUGH 21)*

Variable	Outcome compared to Released/Still Active*	Logit Estimate	Standard Error	t-value	p > |t|
CONSTANT	revoke-arrest	-1.76669	0.1148	-15.39	0.000
	revoke-technical	-2.52492	0.1281	-19.71	0.000
	abscond	-2.36293	0.1226	-19.28	0.000
SEX	revoke-arrest	-0.51044	0.0764	-6.68	0.000
	revoke-technical	-0.25044	0.0661	-3.79	0.000
	abscond	-0.31978	0.0686	-4.66	0.000
LAGEADM	revoke-arrest	-0.38229	0.0272	-14.06	0.000
	revoke-technical	-0.24086	0.0272	-8.84	0.000
	abscond	-0.18360	0.0277	-6.63	0.000
SPLIT	revoke-arrest	0.47928	0.0892	5.37	0.000
	revoke-technical	0.07983	0.1040	0.77	0.443
	abscond	0.43683	0.0915	4.77	0.000
PRPRSN	revoke-arrest	0.42156	0.0352	11.96	0.000
	revoke-technical	0.25642	0.0398	6.44	0.000
	abscond	0.28008	0.0401	6.98	0.000
ADMITS	revoke-arrest	0.15181	0.0199	7.65	0.000
	revoke-technical	0.16653	0.0195	8.54	0.000

Variable	Outcome compared to Released/Still Active*	Logit Estimate	Standard Error	t-value	p > |t|
	abscond	0.11068	0.0204	5.43	0.000
VIOLENT	revoke-arrest	0.13264	0.0686	1.93	0.053
	revoke-technical	0.05580	0.0702	0.79	0.427
	abscond	-0.37351	0.0722	-5.17	0.000
DRUG	revoke-arrest	0.15521	0.0630	2.47	0.014
	revoke-technical	0.32713	0.0590	5.54	0.000
	abscond	-0.09120	0.0613	-1.49	0.137
OTHER	revoke-arrest	-0.05742	0.0951	-0.60	0.546
	revoke-technical	-0.11482	0.0959	-1.20	0.231
	abscond	-0.37651	0.0948	-3.97	0.000
LYRSUP2	revoke-arrest	-0.42754	0.0614	-6.96	0.000
	revoke-technical	-0.45350	0.0627	-7.23	0.000
	abscond	-0.28082	0.0605	-4.64	0.000
LCOUNTS2	revoke-arrest	0.17009	0.0436	3.90	0.000
	revoke-technical	0.11833	0.0455	2.60	0.009
	abscond	0.07500	0.0445	1.69	0.092
CATSUP	revoke-arrest	0.11867	0.0800	1.48	0.138
	revoke-technical	0.41341	0.0767	5.39	0.000
	abscond	0.43712	0.0802	5.45	0.000
CIRCT2	revoke-arrest	-0.54986	0.1864	-2.95	0.003
	revoke-technical	-0.06973	0.2107	-0.33	0.741
	abscond	-0.16780	0.1447	-1.16	0.246
CIRCT3	revoke-arrest	0.23612	0.2661	0.89	0.000

Variable	Outcome compared to Released/Still Active*	Logit Estimate	Standard Error	t-value	p > |t|
	revoke-technical	-0.48236	0.2398	-2.01	0.044
	abscond	0.15258	0.2812	0.54	0.587
CIRCT5	revoke-arrest	0.02182	0.1371	0.16	0.874
	revoke-technical	-0.33436	0.1302	-2.57	0.010
	abscond	-0.06975	0.1352	-0.52	0.606
CIRCT7	revoke-arrest	0.61233	0.2152	2.84	0.004
	revoke-technical	-0.82043	0.2058	-3.99	0.000
	abscond	1.07437	0.1980	5.43	0.000
CIRCT8	revoke-arrest	0.58593	0.2500	2.34	0.019
	revoke-technical	-0.76824	0.2596	-2.96	0.003
	abscond	0.54743	0.2542	2.15	0.031
CIRCT10	revoke-arrest	-0.38440	0.1510	-2.55	0.011
	revoke-technical	0.65877	0.1362	4.84	0.000
	abscond	0.23289	0.1386	1.68	0.093
CIRCT11	revoke-arrest	0.03097	0.1138	0.27	0.785
	revoke-technical	-0.23815	0.1137	-2.09	0.036
	abscond	0.92325	0.1533	6.02	0.000
CIRCT12	revoke-arrest	0.04970	0.1383	0.36	0.719
	revoke-technical	0.29961	0.1506	1.99	0.047
	abscond	0.10133	0.1496	0.68	0.498
CIRCT13	revoke-arrest	-0.62517	0.1303	-4.80	0.000
	revoke-technical	0.40564	0.1222	3.32	0.001
	abscond	0.24546	0.1163	2.11	0.035

Variable	Outcome compared to Released/Still Active*	Logit Estimate	Standard Error	t-value	p > |t|
CIRCT14	revoke-arrest	-0.06302	0.1882	-0.33	0.738
	revoke-technical	0.24804	0.2246	1.10	0.269
	abscond	-0.06346	0.1611	-0.39	0.694
CIRCT15	revoke-arrest	0.27223	0.1306	2.08	0.037
	revoke-technical	0.14412	0.1247	1.16	0.248
	abscond	0.23931	0.2186	1.09	0.274
CIRCT16	revoke-arrest	-0.01853	0.2939	-0.06	0.950
	revoke-technical	0.46398	0.2201	2.11	0.035
	abscond	1.85905	0.2285	8.14	0.000
CIRCT18	revoke-arrest	-0.20802	0.1410	-1.48	0.140
	revoke-technical	-0.70791	0.1367	-5.18	0.000
	abscond	-0.08764	0.1289	-0.68	0.497
CIRCT19	revoke-arrest	0.24185	0.1512	1.60	0.110
	revoke-technical	0.15235	0.1499	1.02	0.309
	abscond	1.66947	0.1654	10.09	0.000
CIRCT20	revoke-arrest	-0.30124	0.1547	-1.95	0.052
	revoke-technical	0.08426	0.1672	0.50	0.614
	abscond	0.18900	0.1475	1.28	0.200
REGION1	revoke-arrest	-0.13318	0.1316	-1.01	0.312
	revoke-technical	-0.22129	0.1646	-1.34	0.179
	abscond	0.48715	0.1244	3.92	0.000
REGION2	revoke-arrest	-0.84819	0.1877	-4.52	0.000
	revoke-technical	0.40047	0.1434	2.79	0.005

Variable	Outcome compared to Released/Still Active*	Logit Estimate	Standard Error	t-value	p > |t|
	abscond	-0.63922	0.1869	-3.42	0.001
REGION3	revoke-arrest	-0.15923	0.1132	-1.41	0.159
	revoke-technical	0.68707	0.1145	6.00	0.000
	abscond	0.20606	0.1158	1.78	0.075
REGION4	revoke-arrest	-0.31265	0.1016	-3.08	0.002
	revoke-technical	0.21931	0.1098	2.00	0.046
	abscond	-1.17707	0.1427	-8.25	0.000

Note: N = 37789 cases

We see in Table 2 that the regional and circuit variables have considerable power in "explaining" differences in the risk of probation failure among otherwise similar subjects. The coefficients of these jurisdictional variables are, most likely, capturing the effects of differences in "local" system philosophy with regard to revocations and absconder warrants, differences in supervision resources and unmeasured differences among regions and circuits in the class of offenders for whom probation is deemed an appropriate sentence.

To give a better understanding of the implications of this model, we examine in the following example the question of the influence of sentence length on failure risk during a specified time interval. Suppose we consider the ratio of failure odds of two probationers who are identical on all variables except that subject A is serving an eighteen-month sentence (LYRSUP2 = 0.405; CATSUP = 1) and subject B is serving a two-year sentence (LYRSUP2 = 0.693; CATSUP = 0). The ratio of their odds of failure during months 16-through-21 (probationer A / probationer B) are:

Revocation for rearrest: 1.3

Revocation for technical violation: 1.7

Abscond: 1.7.

These results are to be interpreted as saying that during this six-month interval A's odds of revocation for rearrest are about 30% higher than B's and his odds of revocation for a technical violation of failure through absconding are about 70% higher.

These results follow directly from equation (1) above. Suppose we denote subject m's odds of failure by mode k relative to non-failure (outcome K) by

$$\Omega_{mk} = \frac{p(k\,|X_m)}{p(K\,|X_m)} = e^{\sum_{j=1}^{J} x_{mj} b_j^{(k)}}. \tag{2}$$

For example, suppose that when equation (2) is evaluated for probationer m, the result has the value 2. This is to be interpreted as meaning that in the course of this interval he is twice as likely to fail by mode k (revocation for rearrest, perhaps) as he is either to be released in the course of the interval or to be still under supervision at its end. Then, from equation (2), the ratio of the k-mode failure odds of probationer A to probationer B are

$$\frac{\Omega_{Ak}}{\Omega_{Bk}} = e^{\sum_{j=1}^{J} (x_{Aj} - x_{Bj}) b_j^{(k)}} \tag{3}$$

In the example, all variable differences between probationer 1 and probationer 2 are 0 except for LYRSUP2 and CATSUP. The results shown above for the failure odds ratios of subject A to subject B follow immediately from the use in equation (3) of the coefficients of these two variables.[14]

[14]Specifically, as noted above, the two subjects differ only on two variables. The difference in the sentence length variable is the difference between ln 1.5 and ln 2 or -0.2877; the difference in the CATSUP variable is 1. The coefficients for the logit comparison of revoke-arrest to released/still active for LYRSUP2 and CATSUP are, from Table 2, -0.42754 and 0.11867, respectively. The arrest risk posed by A compared to B is then calculated as exp(-0.2877*-0.42754 + 1*0.11867), which equals 1.3.

These results might be compared with the ratio of probationer B's failure odds to those of a third subject, C, again identical on all variables except that his sentence length is 32 months (LYRSUP2 = 0.982; CATSUP = 0). Here, the odds ratio of equation (5) depends only on the difference in the log of sentence lengths [LYRSUP2] or ln (2.0) - ln (2.67) = ln (2.0/2.67) = ln (0.75). The odds ratios for these two probationers (probationer B / probationer C) are:

Revocation for rearrest: 1.1

Revocation for technical violation: 1.1

Abscond: 1.1.

What is reflected in this second example is the model's estimate that the failure odds over this six-month interval decrease rather slowly with increasing sentence length. This is not an unexpected result if we can assume that there is on average some measure of rational behavior on the part of the probationer and of the criminal justice system. For, if probation has any deterrent effect, a subject with a long sentence has more to lose through revocation than does a subject sentenced to a relatively short term of supervision. By the same token, if correctional system cost savings are a consideration in the decision to grant probation rather than impose a prison sentence, probation officials and the courts may be more hesitant to revoke offenders under longer sentences. The same dependence on sentence length is, of course, also captured in the first example; but there, in addition, the CATSUP term expresses the tendency of the system to "fail" probationers at the end of their assigned sentence. (See Figures 4, 5, and 6 above.)

2.4 GOODNESS OF FIT

Table 3 shows the standard "Measures of Fit" for the model given in Table 2. These are measures based on the likelihood-ratio test. The "overall" comparison is with a "naive" model--one in which all subjects are assumed to have the same failure probabilities. Here the p-value is a test of the hypothesis that the observed pattern of failures is random--that is, that the set of explanatory variables has no statistical power to "explain" probation failures during this six-month interval. Clearly, this hypothesis can be rejected with confidence. Additionally, as can be seen, the likelihood-ratio tests for individual variables suggest that, with the exception of a few circuit indicator variables, all of the independent variables contribute significantly to the fit of the model.[15]

In all essential characteristics, qualitatively similar comments could be made and similar conclusions drawn based on the parameter estimation results of the models for the other seven time intervals. These results are given in Appendix C.

TABLE 3. MEASURES OF FIT: MODEL 4 (MONTHS 16 THROUGH 21)

Test	-2Log-likelihood Ratio	df	p-value
OVERALL	1941.2644	90	0
CONSTANT	889.4822	3	0
SEX	74.6352	3	0
SPLIT	46.9442	3	0
PRPRSN	187.5929	3	0
ADMITS	129.8088	3	0
VIOLENT	32.715	3	0
DRUG	38.9784	3	0
OTHER	16.8694	3	0.001

[15] The t-statistics associated with individual coefficient values indicate relatively large uncertainties in the estimated values of about one third of the parameters. If the purpose of the investigation were the testing of a theory of probation failure, one would have to consider the hypothesis that the "true" value of these parameters is 0--that certain variables have no influence in determining particular failure modes.. However, in the search for patterns of failure to be found in the data, the estimates reported here were accepted as "best" values, even though it is recognized that the fit between observed and modeled outcomes would not be much degraded by varying the values of low t-statistic parameters over a fairly wide range.

Test	-2Log-likelihood Ratio	df	p-value
CIRCT2	0.6212	3	0.022
CIRCT3	5.4249	3	0.143
CIRCT5	6.8412	3	0.077
CIRCT7	55.3534	3	0
CIRCT8	19.7993	3	0
CIRCT10	33.9455	3	0
CIRCT11	41.9847	3	0
CIRCT12	4.2762	3	0.233
CIRCT13	41.0029	3	0
CIRCT14	1.5727	3	0.666
CIRCT15	6.2845	3	0.099
CIRCT16	68.6559	3	0
CIRCT18	28.2729	3	0
CIRCT19	102.9779	3	0
CIRCT20	6.0766	3	0.108
LAGEADM	288.7846	3	0
REGION1	19.301	3	0
REGION2	41.3672	3	0
REGION3	41.6022	3	0
REGION4	81.6116	3	0
LYRSUP2	109.7991	3	0
LCOUNTS2	22.2658	3	0
CATSUP	55.8309	3	0

Notes:
-2Log-Likelihood for full model: *39857.3605;*
-2Log-Likelihood for restricted model: *41798.6249*
Percent Correctly Predicted: *86.1917*

For a somewhat different investigation of how well individual failure probabilities correspond to observed outcomes, we calculated these probabilities for each subject in each model's construction sample. Subjects were then grouped by incremental intervals of 0.05 in probability for each of the three failure modes. Within each probability interval the expected number of failures by mode k, $<n_k>$, is simply

$$<n_k> = \sum_i p_i^{(k)} \tag{4}$$

where $p_i^{(k)}$ is the probability that subject *i* will fail by mode *k* and the sum is over all subjects active at the beginning of that interval. We can also calculate the standard deviation of $<n_k>$ as

$$SD(<n_k>) = \sqrt{\sum_i p_i^{(k)}(1 - p_i^{(k)})} \tag{5}$$

The observed and expected numbers of failures in the population at risk over the period from month 16 through month 21 are shown in Figures 8 through 10 for the three failure modes--revoke for a new arrest, revoke for a technical violation, and abscond, respectively. The two standard deviation band is also shown. Similar figures are given for other models in Appendix C. As can be seen, there are few deviations between observed and expected numbers of failures that exceed the two standard-deviation bands.

Thus, there would not appear to be any statistical reason for rejecting the failure probability distributions defined by these models. Some degradation in "predictive power" must be anticipated when the models are applied to samples other than the construction sample. But unless experience demonstrates otherwise, the model-assigned, individual failure probabilities seem to offer a reliable basis for a failure risk classification of probationers under active supervision.

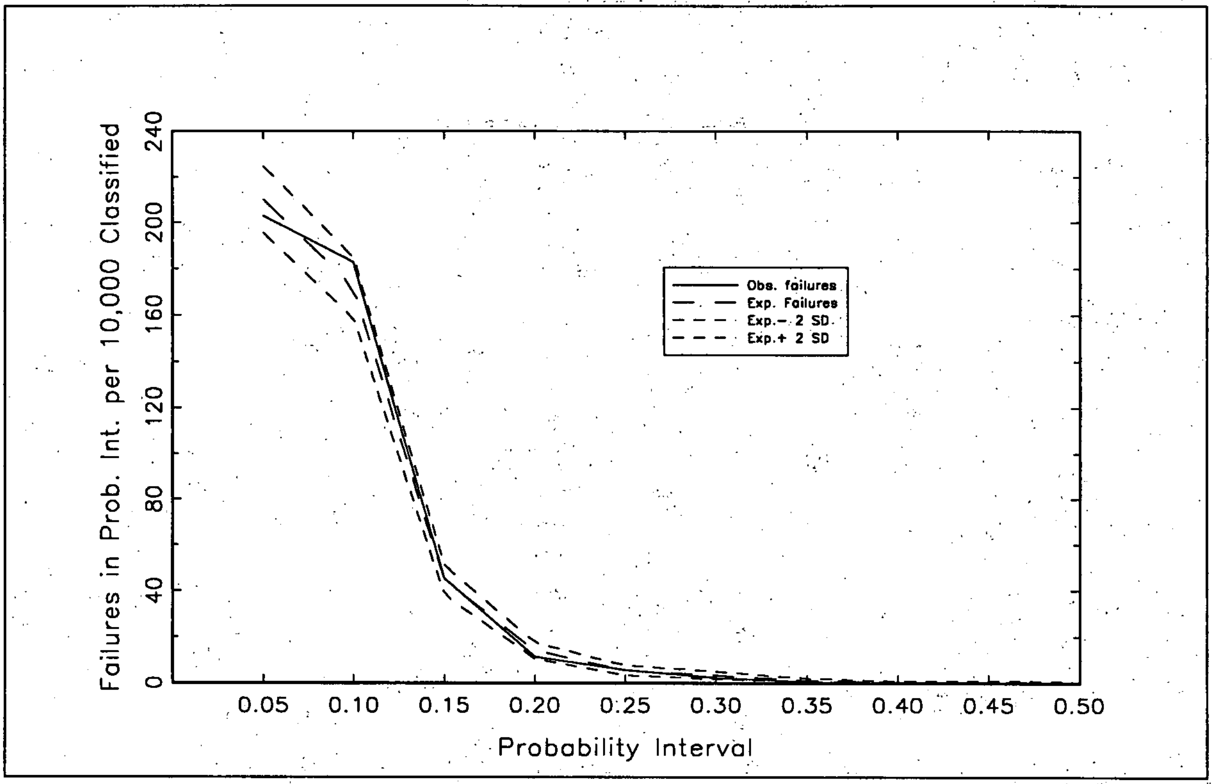


Figure 8. Goodness of Fit for Model 4 (Months 16 -21), Revocation for New Arrest.

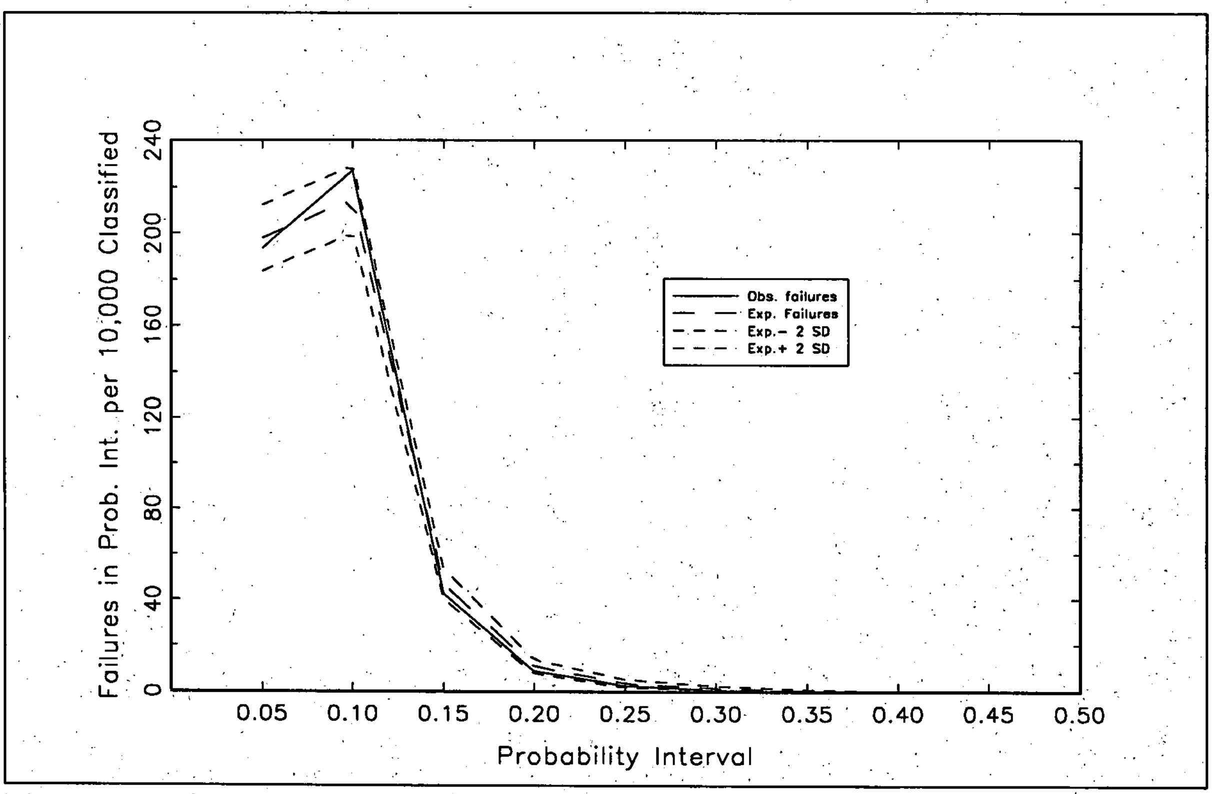


Figure 9. Goodness of Fit for Model 4 (Months 16-21), Revocation for Technical Violation.

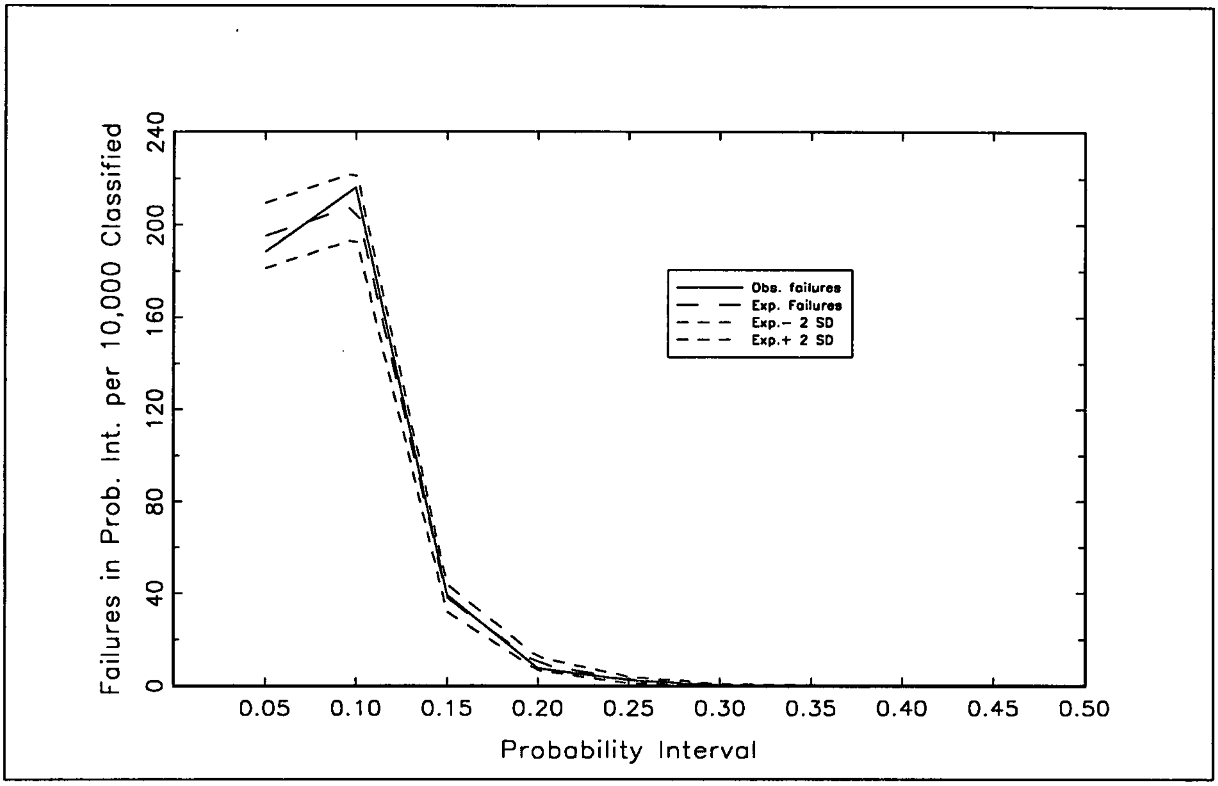


Figure 10. Goodness of Fit for Model 4 (Months 16-21) for Absconding.

3.0 CURRENT AND FUTURE ACTIVITIES

This section outlines the tasks underway or planned by the Florida Department of Corrections' Bureau of Planning, Research and Statistics in order to convert the technical results reported in this paper into an automated operational system for a triage classification of active probation cases.[16]

One of the first steps in this process is the policy determination of the population of cases that will be subject to classification based on a statistical risk assessment. In particular, the Department of Corrections plans to override the statistical assessment for certain classes of probationers for whom policy dictates a maximum supervision level throughout their sentence --perhaps because of the nature or severity of the conviction offense. It was decided that the following types of cases will be *automatically classified as high risk:*

- Conditional release offenders
- Conditional medical release offenders
- Lewd and lascivious offenders
- Sexual battery offenders
- Child abuse offenders
- Sexual predator offenders
- Board of clemency offenders
- Habitual and violent offenders (designated by Courts)
- Sentencing guidelines Level 8 and above offenders.

The next step is to examine the actual distribution over failure outcomes of an active probation caseload. In late September 1995, a distribution of six-month failure risk

---

[16]Findings reported in this section are based on information and results kindly supplied by Kristine Leininger, Florida Department of Corrections, Bureau of Planning, Research and Statistics.

probabilities was generated for the universe of over 78,000 active cases that were classifiable by policy and that had complete records on all the models' independent variables. In these calculations the parameter values of the model covering months 4 through 9 were used to estimate failure probabilities for all probationers who had not yet completed nine months under supervision. For subjects at risk for more than nine months but less than forty months, the model was used that corresponded to the length of time since they had been admitted to probation. For all cases active more than 39 months, probabilities were based on the model estimated for the final time interval--months 40 through 48. This model differs from the others in that it estimates nine- month failure probabilities. For purposes of defining a classification level, these nine-month probabilities were multiplied by 0.667, thus assuming that the probability of failure is uniformly distributed over this interval.

The failure probability distributions over active probation cases (as of September 1995) are shown in Figures 11 and 12. About 5% of the currently active population have a six- month probability of failure by any mode of 0.04 or less; and about 5% have a probability of 0.30 or greater. The mid-point of the distribution occurs at a probability of 0.13 or 0.14. These results or similar jurisdiction-specific distributions of failure probabilities will serve as a basis for specifying the probability ranges to be classified as low, medium or high risk.

It is of some interest to examine how failure probabilities are distributed over the seven time intervals from new admissions through long-term survivors active for more than 40 months. As might be expected, the numbers of subjects under active supervision decrease over successive intervals:

Less than 9 months	N =	31, 346
10 - 15 months		14,459
16 - 21 months		9,486
22 - 27 months		6,613
28 - 33 months		4,445
34 - 39 months		3,295
Greater than 39 months		8,558.

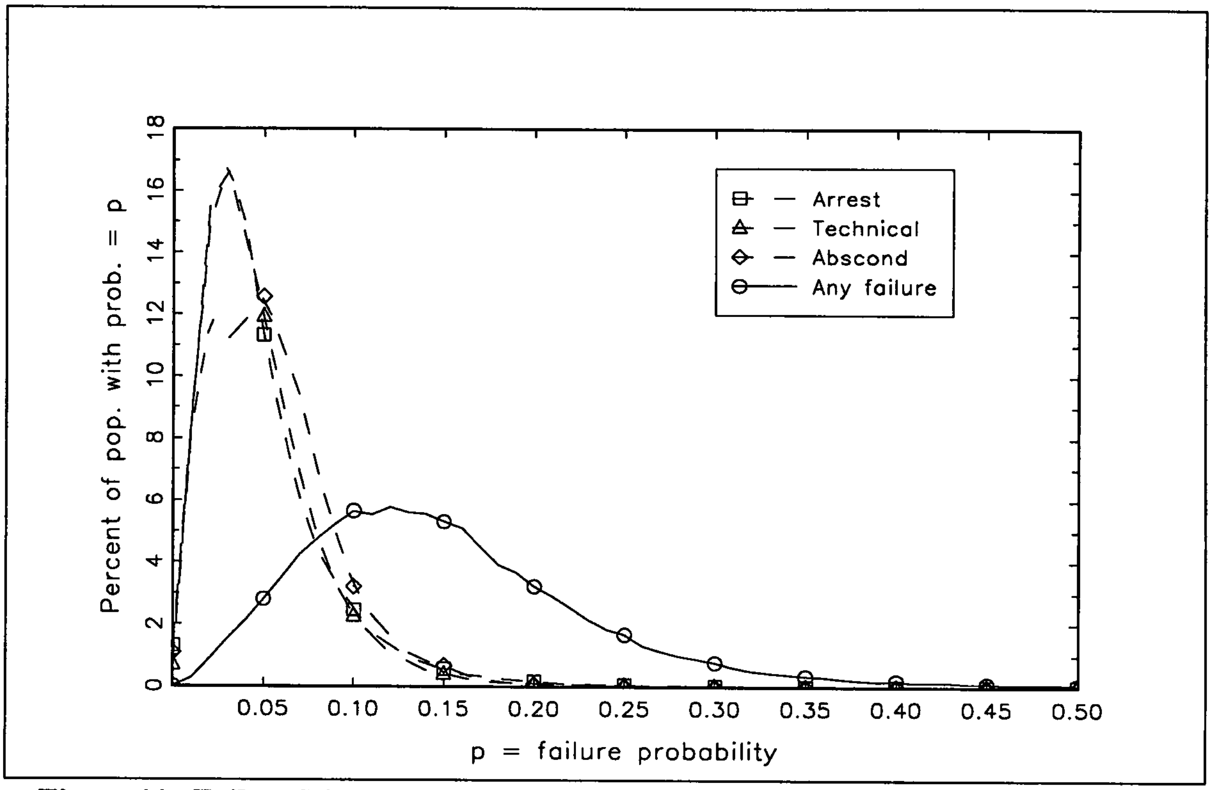


Figure 11. Failure Distribution for Florida Probation Population (circa September 1995).

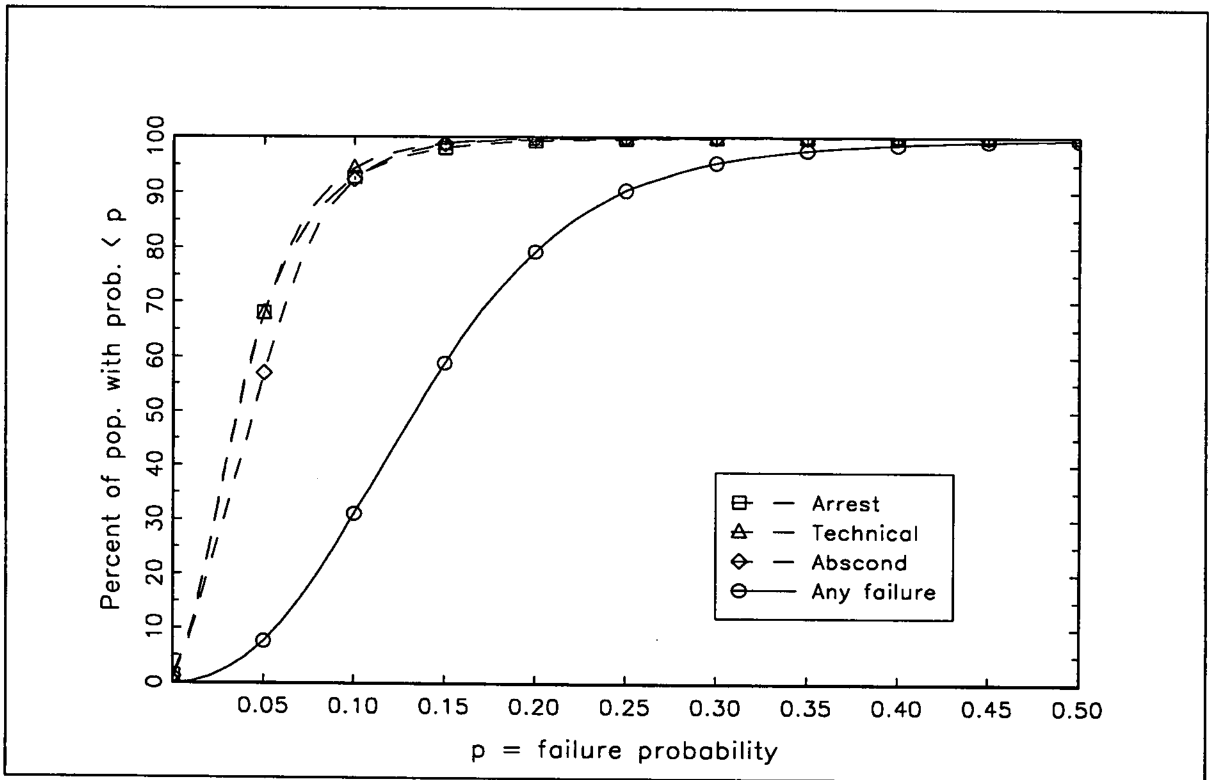


Figure 12. Cumulative Failure Distribution for Probation Population (circa September 1995).

In great part, of course, this decrease is due to probationers being released from supervision after successful completion of their sentence. But in some part it also reflects losses to the active population through revocation orders or absconder warrants. Figures 13 and 14 show the considerable differences in the distribution of failure probabilities for populations making up three of the seven time intervals:

- Those with less than nine months under supervision,
- Those who have served more than twenty one months but less than twenty eight months of their sentence, and
- Those who have been active for more than 39 months.

What is readily seen is the systematic decrease as time goes on in the overall failure risk of population surviving in active supervision status. This suggests that individual failure risks may, in general, also be decreasing over time as the average failure probability for survivors decreases over successive intervals.

The probabilities on which these distributions are based were calculated using a mainframe program written in SPSS. The plan is to translate this code into the language used by the management information system of offender case files. Probabilities will be automatically recalculated monthly for all cases under active supervision and appropriate risk classifications assigned. The information on individual failure probabilities and risk classifications will then be furnished to the field offices responsible for case supervision.

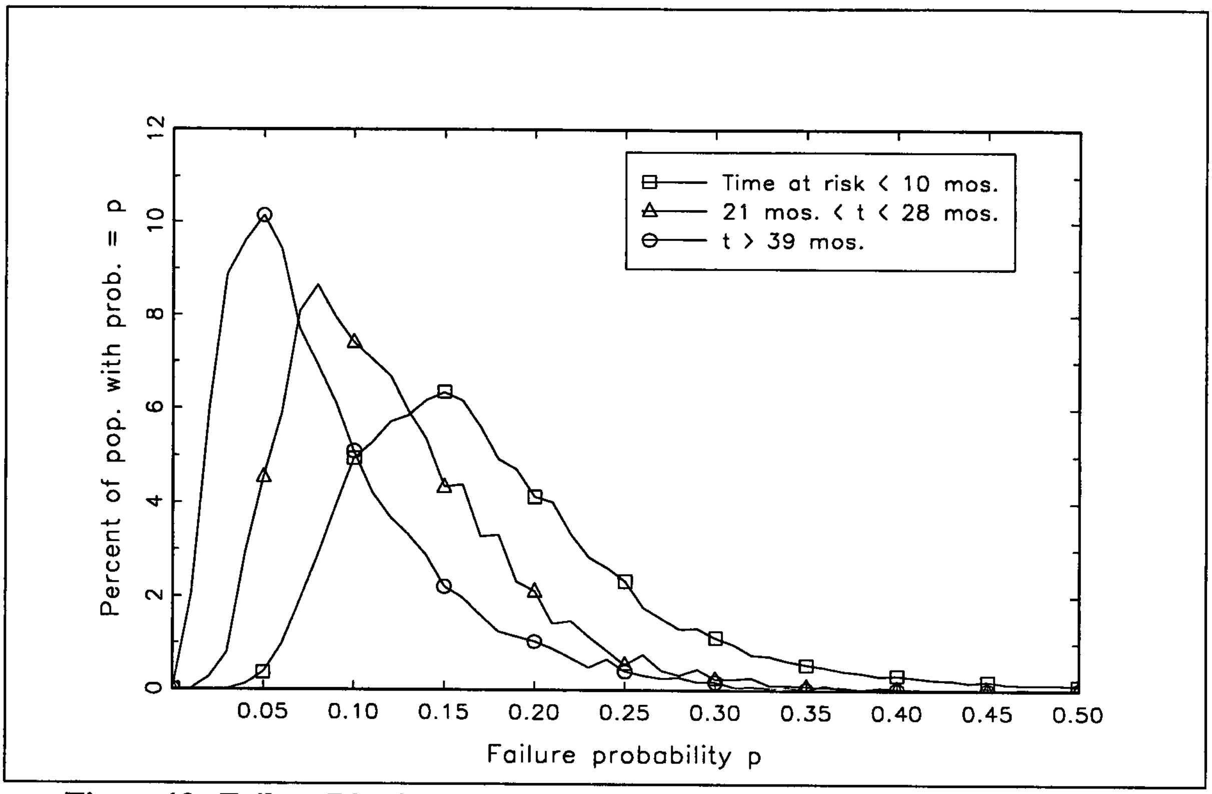


Figure 13. Failure Distributions for Three Periods on Probation, Florida Data.

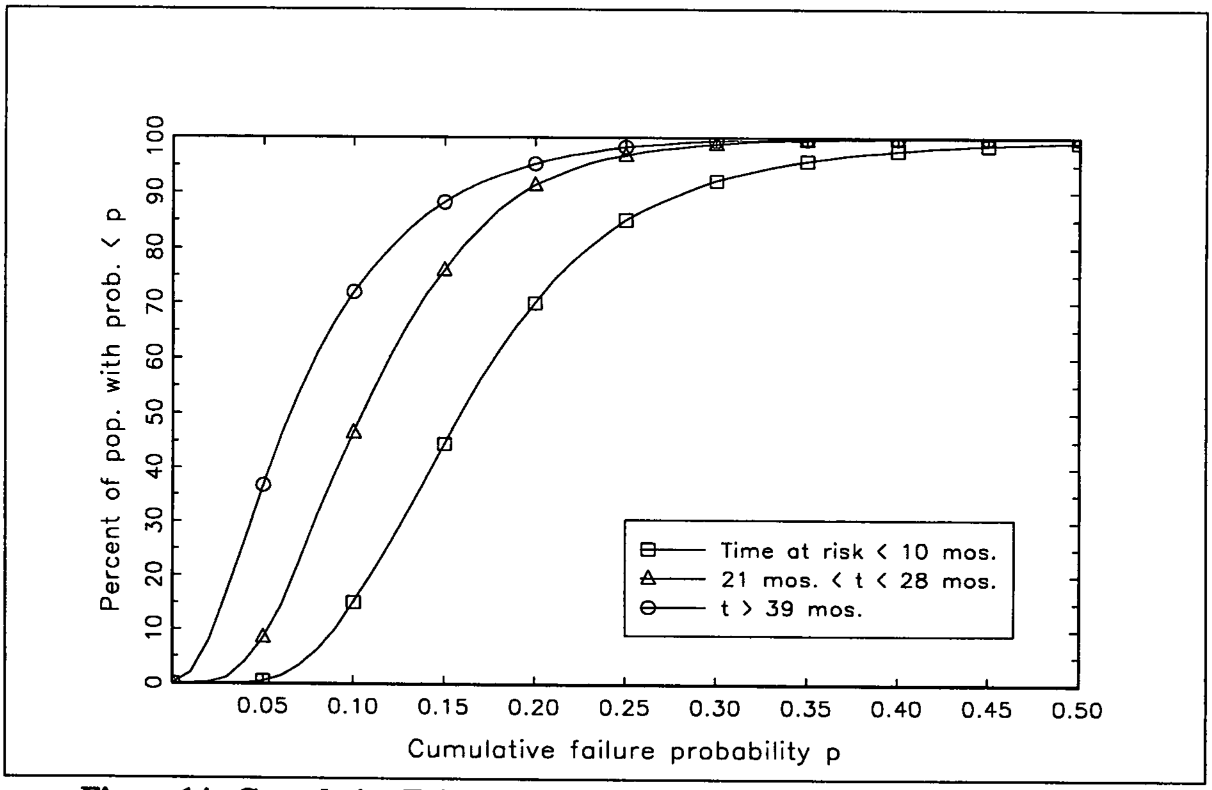


Figure 14. Cumulative Failure Distributions for Three Periods of Probation.

Another use that can be made of these probabilities is an estimation of the numbers and rates of failures expected over the next month. If we assume that the failure probability densities are approximately uniformly distributed over each interval, then $p_i/6$ is an approximation to the one-month ahead failure probability of subject I, where p_i is the probability assessed by the model for a six month interval. Substituting $p_i/6$ for p_i in Equations (4) and (5) of Section 2, we obtain a one-month approximation to the expected numbers of failures and their standard deviations. The results are shown in Table 4, both in terms of expected numbers of failures based on the population under active supervision at the time the probability distributions were generated and in terms of the expected failure rates. As can be seen, of the approximately 78,000 cases, 1,944 or 2.5% are expected to fail in the next month.

TABLE 4. EXPECTED ONE-MONTH FAILURE PROJECTIONS

Failure Mode	Expected Failures	2 Std. Dev. Range of Failures	Expected Fail Rate	2 Std. Dev. Range of Rate
Revoke Arrest	644	594 - 695	0.82%	0.76% - 0.88%
Revoke Technical	624	574 - 674	0.80%	0.74% - 0.86%
Abscond	705	652 - 758	0.90%	0.83% - 0.97%
Any Failure	1944	1857 - 2031	2.50%	2.39% - 2.61%

Table 4 gives estimates for the state as a whole. Similar results could be generated for any well defined sub-population of reasonable size simply by restricting the sums in equations (4) and (5) to members of that sub-population.[17] For example, there might be an interest in calculating expected one-month failure rates for particular regions or circuits.

Projections such as these have practical importance. First, of course, this information can help managers allocate resources, since each failure will place time demands on probation officers to prepare paperwork, present revocation petitions in court, etc. Secondly, however, a succession of actual monthly fail rates that lie outside the two standard deviation band would give an indication that the process by which failures are generated may have changed significantly from what was found in the failure patterns of the 1991-1994 data. There are a number of reasons why such a change might occur. They might, for example, indicate unmeasured changes in the character of the more recent admissions cohorts or pragmatic changes in the pragmatic policies of probation officers and courts with regard to failure decisions. Perhaps most interesting from the perspective of probation supervision would be changes that might be attributable to implementation of a new policy on intensity of supervision. Monitoring of the expected vs. observed fail rates could, thus, provide an early warning signal that something of interest is going on that is deserving of closer investigation. In any case it could signal a need for re-estimation of the probability models with more recent data.

[17]For very small sub-populations the results would not have much meaning. This is because the ratio of the standard deviation of expected number of failures to the expected value varies as $1/\sqrt{N}$, where N is the sub-population size. Thus, for small N, the 2 standard deviation band would be very wide relative to the number of failures expected.

APPENDIX A

An Overview of Logistic Regression Models:

Definitions

Subscripts:

Let

i identify a particular subject (i = 1, 2, ... N);

j identify a particular independent variable including the intercept term (j = 1, 2, ... J);

k identify a particular value of the dependent (i.e. the outcome) variable (k = 1, 2, ... K).

We assume that the K outcomes are mutually exclusive and exhaustive. Every case is observed to have one and only one of these outcomes.

Further, let

$\mathbf{X}_{ij}$ be the (N x J) data matrix of explanatory variables (including an intercept term);

$\boldsymbol{\beta}_{jk}$ be the (J x K-1) matrix of model coefficients;

$\mathbf{p}_{ik}$ be the (N x K) matrix of outcome probabilities.

and

$$D_i = 1 + \sum_{k=1}^{K-1} e^{\sum_J x_{ij} \beta_{jk}}. \qquad \textbf{(A-1)}$$

Then, under a multinomial logit model,

$$p_{ik} = \frac{e^{\sum_J x_{ij} \beta_{jk}}}{D_i}. \qquad \textbf{(A-2)}$$

for k = 1, 2, ... K-1; and

$$p_{iK} = \frac{1}{D_i}. \tag{A-3}$$

To see the derivation of the model, we assume that

$$\ln \frac{p_{ik}}{p_{iK}} = \sum_j X_{ij} \beta_{jk}. \tag{A-4}$$

This is the model's distributional assumption: the log of the odds of subject I having outcome k relative to outcome K is a linear function of the explanatory variables **X**. Then

$$p_{ik} = p_{iK} \, e^{\sum_j X_{ij} \beta_{jk}} \tag{A-5}$$

Summing these equations from k = 1 to K-1 and adding p_K to both sides, we get

$$p_{iK} + \sum_{k=1}^{K-1} p_{ik} = p_{iK} D_i. \tag{A-6}$$

But since the K outcomes are exhaustive and mutually exclusive, the left hand side is equal to 1 and the equations for p_{iK} and the p_{ik} follow immediately.

The Likelihood Function and the Maximization Equations

Suppose outcome k(I) is observed for subject I. His contribution to the log of the likelihood function is then

$$\ln p_{i\,k(i)} = \sum_j X_{ij} \beta_{j\,k(i)} - \ln D_i \tag{A-7}$$

for k(I) one of the outcomes 1, 2, ... K-1. For k(I) = K, it is simply

$$\ln p_{iK} = -\ln D_i. \tag{A-8}$$

Thus, in a rather clumsy but straightforward notation the log likelihood becomes

$$\ln L = \sum_{i,\,k(i)=1} \sum_{j} X_{ij} \beta_{j1} + \sum_{i,\,k(i)=2} \sum_{j} X_{ij} \beta_{j2} + \ldots - \sum_{all\ i} \ln D_i \tag{A-9}$$

The set of double sums runs from k(I) = 1 to k(I) = K-1. Each, of course, is a sum only over those subjects observed to have outcome k(I). The sum over all subjects follows from the fact that all outcome probabilities, k = 1 to k = K, have the denominator D.

The first derivative with respect to a particular coefficient $\beta_{j1\,k1}$ (i.e. j = j1; k = k1) is then

$$\frac{\partial \ln L}{\partial \beta_{j1\,k1}} = \sum_{i,\,k(i)=k1} X_{ij1} - \sum_{all} \frac{\partial \ln D_i}{\partial \beta_{j1\,k1}} = \Phi_{j1,\,k1} \tag{A-10}$$

where

$$\frac{\partial \ln D_i}{\partial \beta_{j1\,k1}} = \frac{1}{D_i} X_{ij1}\, e^{\sum_j X_{ij} \beta_{jk1}}. \tag{A-11}$$

Again note that the first sum on the right in the equation for the derivative of ln L runs only over those subjects observed to have outcome k1.

The maximization problem is to find solutions for the system of J x K-1 equations

$$\Phi_{j1,\,k1} = 0 \tag{A-12}$$

for values j1 = 1, 2, ... J and k1 = 1, 2, ... K-1.

There are three things that might be noted about this system of equations. First, each of the J x K-1 equations involves all of the coefficients. The problem cannot be partitioned into sub-systems of smaller dimensions. What this means is that, if for some reason a solution for one of the coefficients doesn't exist, then no solution exists for any of them.

Second, substitution of equations (A-2), (A-10), and (A-11) into equation (A-12) gives

$$\sum_{i,\, k(i) = k1} X_{i\,j1} = \sum_{all} X_{i\,j1}\, p_{i\,k1}\,. \qquad \textbf{(A-13)}$$

Suppose j1 is the intercept term for outcome k1 so that $X_{i\,j1} = 1$ for all I. The sum on the left of (A-13) is then simply the number of subjects <u>observed</u> to have outcome k1 and the sum on the right is the <u>expected</u> number of k1 outcomes. The maximization equations thus guarantee that, for the population as a whole, the expected and observed numbers will be identical for each outcome. Obviously, this will also be true if the variable $X_{i\,j1}$ is categorical with possible outcome values of 0 and 1. Thus, it is no surprise that, for example, the observed number of females (or males) to have any particular outcome is exactly equal to the expected number "predicted" by the model.

However, an exception of practical importance occurs when all observed outcomes of type k1 are associated with only one of the values of a dichotomous variable--only males, perhaps. In that case, equation (A-13) would necessarily be inconsistent with the data. For example, suppose all subjects with outcome k1 were defined as having $X_{i\,j1} = 0$. The left hand side of equation (A-13) then is 0 while the right hand side, the sum over all subjects, is necessarily positive. In that case, the routine for estimating the model's parameters will simply fail to converge. This situation is not unusual in a model with many dichotomous independent variables. Should this occur, one possible solution is to combine two or more independent variables into a single new one--if that makes sense theoretically. For example, in one of the models estimated for this study, it was convenient to define a new variable as a combination of two small judicial circuits. Cases could arise however in which this kind of solution is simply nonsense. For example, suppose there were no female subjects revoked for rearrest. Gender would then seem to play a quite strong, non-ignorable role. The simplest solution in that case would seem to be to model male and female outcomes separately.

APPENDIX A
SOME MATHEMATICAL RELATIONS USED IN THE STUDY

Definitions and Comparisons of the Negative Binomial and Poisson Distributions

Suppose that during a time interval of length t_i in which he was free in society and, hence, at risk of arrest given an offense, subject i is arrested y_i times. Under the *a priori* assumption of a negative binomial distribution for the numbers of arrests in this time interval, the probability of exactly y_i arrests is given by

$$p_{nb}(y_i \mid \lambda_i,\ \sigma_i^2,\ t_i) = \frac{\Gamma\left(\dfrac{\lambda_i t_i}{\sigma_i^2 - 1} + y_i\right)}{\Gamma\left(\dfrac{\lambda_i t_i}{\sigma_i^2 - 1}\right) y_i!} \left(\frac{\sigma_i^2 - 1}{\sigma_i^2}\right)^{y_i} (\sigma_i^2)^{-\frac{\lambda_i t_i}{\sigma_i^2 - 1}} \tag{A.1}$$

where $\Gamma(.)$ is the gamma function:

$$\Gamma(\kappa) = \int_0^\infty x^{\kappa - 1} e^{-x}\, dx. \tag{A.2}$$

The parameters of this distribution, λ and σ^2, are assumed to be constant in time but dependent on a set of covariates $\mathbf{X}_i$ that characterize this subject. We assume these parameters to be log linear in the components of $\mathbf{X}$:

$$\lambda_i = e^{X_i \beta} \tag{A.3a}$$

and

$$(\sigma_i^2 - 1) = e^{X_i \gamma} \tag{A.3b}$$

With this parametrization, which follows King (1989), the expected number of arrests in a time interval t for subjects described by the vector **X** is

$$< y(\boldsymbol{X}) > \; = \; \lambda\,(\boldsymbol{X})\,t \tag{A.4}$$

with a variance given by

$$var\,(y\,(\boldsymbol{X})) \; = \; \sigma^2\,(\boldsymbol{X})\,\lambda\,(\boldsymbol{X})\,t. \tag{A.5}$$

From (A.3b) it follows that σ^2 must be greater than 1. In the limit as it approaches 1, the negative binomial distribution for numbers of arrests in time t approaches the Poisson:

$$p_{Poisson}\,(y \mid \lambda, t) \; = \; \frac{(\lambda\, t)^y}{y\,!}\, e^{-\lambda\, t}. \tag{A.6}$$

For a given λ, the Poisson expected value of y is again given by (A.4). The variance of y, however, is simply λt, which is smaller than the negative binomial variance of (A.5). (In the statistical literature this is termed "over dispersion" and σ^2 the "dispersion parameter.") Indeed, one motivation for the development of the negative binomial distribution is the desire to relax the Poisson's rather stringent assumption that the mean and variance of y are necessarily equal.

This "over dispersion" has some simple consequences of relevance in comparing the results of the Poisson and negative binomial distributions with the same frequency parameter λ. In general, the negative binomial is "flatter" than the Poisson, meaning that the probabilities are greater both at the low and high ends of the distribution of numbers of arrests in a given time interval.

Consider the ratio of negative binomial and Poisson probabilities of no arrests in time t. From equations (A.1) and (A.6) this is

$$\frac{p_{nb}(y=0 \mid \lambda, t, \sigma^2)}{p_{Poisson}(y=0 \mid \lambda, t)} = e^{\lambda t\left(1 - \frac{\ln \sigma^2}{\sigma^2 - 1}\right)}. \tag{A.7}$$

It can be shown that for $\sigma^2 > 1$ the right hand side of (A.7) is everywhere greater than 1. It approaches 1 in the limit as σ^2 approaches 1 and has a limiting value of $e^{\lambda t}$ as σ^2 becomes very large.

At the other end of the y distribution, we consider the recurrence relations implied by equations (A.1) and (A.6). For the negative binomial distribution we have

$$p_{nb}(y+1 \mid \lambda, \sigma^2, t) = \frac{\lambda t + (\sigma^2 - 1)y}{\sigma^2 (y+1)} p_{nb}(y \mid \lambda, \sigma^2, t). \tag{A.8a}$$

For the Poisson distribution

$$p_{Poisson}(y+1 \mid \lambda, t) = \frac{\lambda t}{y+1} p_{Poisson}(y \mid \lambda, t). \tag{A.8b}$$

For large values of y, the ratio of successive Poisson probabilities decreases fairly rapidly as $(y+1)^{-1}$. In marked contrast, for large y, successive terms of the negative binomial distribution decrease only as y / (y+1).

The general implications of the dispersion parameter σ^2 for interpretation of model results are straightforward. For a given $\lambda(\mathbf{X})$ the variance in the number of arrests that would be observed in a given time interval t increases with increasing $\sigma^2(\mathbf{X})$. Thus, with a good model we might be able to predict with reasonable accuracy the mean number of arrests expected during time t among a large population that is homogeneous with respect to measured characteristics **X**. But we must anticipate that, as σ^2 increases, individual arrest counts will range quite widely on either side of this mean, making subject-level prediction problematic.

Unobserved Heterogeneity as a Process Underlying the Negative Binomial Distribution

A negative binomial distribution for counts of events can arise under a quite broad set of assumptions regarding the more fundamental stochastic processes that generate the observed events. In the published literature on criminal careers the critical assumption is unobserved heterogeneity.

We assume first that at least for realistically measurable time intervals each subject's offending behavior remains essentially unchanging so that the number of arrests observed per unit time are a realization of a stochastic process described by a Poisson distribution with constant, true rate parameter λ_T. (Unless otherwise noted we will from here on as well as in the body of the report give results in terms of an observation period of t = 1 year of street time. The variable t is suppressed in the following equations.) Thus, for subject i, the probability of observing y arrests in 1 year at risk is

$$p_{Poisson}(y \mid \lambda_{iT}) = e^{-\lambda_{iT}} \frac{(\lambda_{iT})^y}{y!} \tag{A.9}$$

Of course, the expected value and variance of subject i's annual arrest rate are then both equal to $\lambda_{iT,}$ which we now assume is both unknown and unobservable at the level of the individual subject.

Suppose, however, that for all subjects we have measured a set of theoretically relevant covariates **X** and that subject i belongs to a class that is homogeneous with respect to all components of **X**. We now make the further assumption that within this class the true arrest rate parameter λ_T is a random variable that follows a two-parameter gamma distribution with probability distribution function

$$g_\Gamma(\lambda_{iT} \mid \theta(X_i), \kappa(X_i)) = \frac{(\lambda_{iT})^{\kappa - 1} e^{-\frac{\lambda_{iT}}{\theta}}}{\theta^\kappa \, \Gamma(\kappa)}. \tag{A.10}$$

From equations (A.9) and (A.10) it follows that we are assuming that the joint probability that subject i has a true arrest rate parameter in the range $(\lambda_{iT}, \lambda_{iT} + d\lambda_{iT}]$ and would be observed to experience y arrests per year free is

$$p_{Poisson}(y \mid \lambda_{iT})\, g_{\Gamma}(\lambda_{iT} \mid \theta, \kappa)\, d\lambda_{iT}.$$

By assumption, subject i's true rate parameter λ_{iT} cannot be observed but we can determine the marginal distribution of arrest counts per unit time for the class $\mathbf{X}_i$ by integrating the joint probability distribution over all possible values of λ_{iT}. Translating the gamma distribution parameters $\theta(\mathbf{X})$ and $\kappa(\mathbf{X})$ back to the parameters $\lambda(\mathbf{X})$ and $\sigma^2(\mathbf{X})$,

$$\theta(X) = \sigma^2(X) - 1 \tag{A.11a}$$

and

$$\kappa(X) = \frac{\lambda(X)}{\sigma^2(X) - 1}, \tag{A.11b}$$

we obtain the negative binomial distribution in the form of equation (A.1).

From the properties of the gamma distribution (A.10), it follows that the expected value of the true annual arrest rate for the class $\mathbf{X}_i$ is then

$$< \lambda_{iT} \mid X_i > = \lambda(X_i) \tag{A.12a}$$

and its variance is

$$var(\lambda_{iT}) = \lambda(X_i)\,(\sigma^2(X_i) - 1) \tag{A.12b}$$

Note that this is the variance of the unobserved true arrest rate within an **X**-homogeneous population.

It is convenient to define an "unobserved heterogeneity index" $\zeta(\mathbf{X})$ associated with any population class **X**. Using (A.12a) and (A.12b), we can write

$$\zeta(X_i) = \sigma^2 - 1 = var(\lambda_{iT}|X_i)/\lambda(X_i) = var(\lambda_{iT}|X_i)/<\lambda_{iT}|X_i> \quad \text{(A.13a)}$$

and from (A.3b) we have

$$\zeta(X_i) = e^{X_i\gamma}. \quad \text{(A.13b)}$$

Obviously, if ζ is close to 0, a population homogeneous with respect to **X** is also relatively homogeneous with respect to individual arrest rates. Conversely, a large value of ζ would indicate substantial heterogeneity in the true arrest rates, despite the fact that this subset of the offender population is homogeneous with respect to all measured covariates.

Elasticities of λ(X) and ζ(X)

The elasticity of a function at a point **X** with respect to a particular component x_k is defined as the percent increase in the value of the function that would be induced by a 1% increase in X_k, all other component values held constant.

From (A.3a) it follows that the fractional change in λ(X) accompanying a vector of arbitrary but small changes in **X** is

$$\frac{d\lambda(X)}{\lambda(X)} = \beta\, dX \quad \text{(A.14)}$$

Let $dX_k = 0.01\ X_k$ with all other $dX_j = 0$. Multiplying both sides by 100 to express the results in percent, we obtain the simple form of the elasticity of λ with respect to a change in the variable X_k as

$$\epsilon_{\lambda k} = \beta_k X_k \tag{A.15}$$

From (A.3b) and (A.13), a similar result holds for ζ. Specifically, we can define the elasticity of ζ

$$\epsilon_{\zeta k} = \gamma_k X_k \tag{A.16}$$

Note that the elasticity depends on the point **X** at which it is calculated. In this study we examine elasticities at the population means of the covariates. Because of the log linearity of λ and ζ the elasticities are then a measure of the variability of the population geometric means of these functions with respect to small changes in mean values of the covariates.

APPENDIX B

Software for Estimating Multinomial Logit Models

The model parameters reported in this paper were estimated using the Gauss Quantal Response library procedure. Gauss is a matrix-based system for mathematical manipulation of data. It was developed and is marketed by Aptech Systems, Inc. of Maple Valley, Washington. With a data base of about 50,000 observations and a model with 93 parameters, convergence to a tolerance of 10^{-4} was typically achieved in 4 or 5 iterations. Using a Pentium PC, the time required was about 6 minutes. LIMDEP, a statistical package available from Econometric Software, Inc., of Bellport, New York, also has a procedure for estimating logistic regression models with multinomial outcomes. Further, we were informed that these models can be estimated in GENSTAT 5, a statistical software system developed at the Rothamsted Experimental Station in England and distributed in the U.S. by National Algorithms Group, Inc., of Downers Grove, Illinois. Both LIMDEP and GENSTAT are available in either PC or mainframe versions. There are undoubtedly other statistical packages that can carry out these computations. Those mentioned here are simply ones that are known to the authors or that happened to come to their attention.

SPSS and SAS both offer efficient procedures for estimating logistic regression models with a binomial outcome--addressing the question "fail" or "not fail" without making distinctions between different modes of failure. However, the routine for estimation of multinomial logit models is in both cases embedded in the log linear analysis procedure. This method of analysis involves certain difficulties for the kind of application considered here.

The first problem is that all independent variables must be defined as categorical. This creates something of a dilemma for the coding of variables that are, in fact, continuous or integer level (e.g. age at admission, sentence length, number of prior prison commitments). Either the definition of categories requires the placement of cut points in the data for which theory can offer only limited guidance or else the number of parameters to be estimated must quickly grow to an unmanageable size. For example, if age at admission is defined by K Categorical levels and there are three failure outcomes in the model, the age variable will require that 3 x (K-1) parameters be estimated rather than the 3 that would be required if age could be treated as continuous or integer level.

A second problem seems to be that the log linear procedure is not very efficient for estimating logit models that are based on a large population with a relatively large number of independent variables. This is probably due to the amount of computation time required to categorize observations into the contingency table cells on which the analysis is based rather than to the time needed for the likelihood maximization routine used for parameter estimation.

We explored using the SAS CATMOD procedure on a data base of about 50,000 observations with a total of 120 parameters to be estimated. More than 11 minutes of mainframe CPU time were needed to achieve convergence. With the same data and variable set, convergence was achieved on a PC using the Gauss logistic regression routine in about 6 minutes.

APPENDIX B

FINAL MODEL FORMS USED IN ANALYSES

The coefficients of the "final," relatively parsimonious models are shown in the two following tables. These forms were arrived at by successively deleting variables having low values of the t-statistics in previous estimation runs. The criterion used for retention in the final models was the probability p(t) less than or equal to 0.10.[14]

The analysis data base includes 42 variables. We initially estimated a model with a rate function $\lambda(X)$ that included an intercept term β_0 and coefficients for all variables and a variance function ζ limited to an intercept term γ_0. We then estimated a model with all variables included in both the rate and variance functions--a total of 86 parameters. Likelihood-ratio tests were used to investigate whether significant information was generated by including covariates in the variance function. The test statistic for the model for Year 1 is $\chi^2 = 95.41$ with 42 degrees of freedom (p-value = 5×10^{-6}). In the model for Years 2 and 3, the corresponding χ^2 test statistic and p-value are 70.3 and 0.004, respectively. We can clearly reject the hypothesis that covariates in the variance function contribute no significant information.

We also calculated likelihood-ratio test statistics to compare the full, 86-parameter model to the "final" forms shown in the Tables B.1 and B.2. The test statistic for Model Year 1 is $\chi^2 = 20.6$, df = 43, and $p(\chi^2) = 0.9985$. For Model Years 2 & 3, $\chi^2 = 26.4$, df = 46, and $p(\chi^2) = 0.991$. Apparently, setting "nonsignificant" model parameters equal to 0 (as determined by t-statistics) has not resulted in any statistically significant loss of information.

Some readers may find it curious that the log likelihood is positive for the fitted model covering Years 2 and 3. The parameters were estimated using the negative binomial option in the GAUSS™ application for count models (King, 1995). In this procedure, the intrinsically non-positive term $-\sum \ln y_i!$ (See equation A.1) is simply dropped from the

[14]In the final model for Years 2 and 3 post-release, the intercept coefficient of the variance function γ_0 was retained although not statistically significant at 0.10.

calculation of the log likelihood value. Upon differentiation with respect to the parameters, this term would disappear and, thus, it plays no role in model estimation.

Table B.1. Negative Binomial Model Results (First Year Post-Release)

Variable	Parameter Estimate	S.E.	Heteroskedastic-consistent S.E.
β			
Intercept (β_0)	-0.6488	0.3915	0.387
COHORT	-0.1326	0.0788	0.0792
BAYAREA	0.2413	0.059	0.0584
SONOTLA	0.2038	0.0714	0.0695
HISPANIC	0.1834	0.057	0.0601
BLACK	0.3631	0.0617	0.0638
AGEREL	0.0244	0.0164	0.0165
DISCHGED	0.271	0.0825	0.0807
YAGANG	0.1022	0.0348	0.0361
YAVIO	0.081	0.0298	0.0303
INFRATE	0.073	0.0147	0.0152
FIRSTADM	-0.469	0.092	0.0971
ROBBER	0.101	0.0288	0.0282
BURGLAR	0.063	0.0142	0.0136
OTHPRO	0.0672	0.0144	0.0148
MISCHG	0.0522	0.0065	0.007
PREVIO	-0.0768	0.0272	0.0284
ALCOHOL	-0.046	0.0331	0.0341
DRUGABU	0.098	0.0334	0.0333
PARTNSAD	-0.171	0.0674	0.073
DROPOUT	0.1232	0.0423	0.0452
CONTROL	0.0493	0.0279	0.0294
SCHDISC	-0.0394	0.0277	0.0288

Variable	Parameter Estimate	S.E.	Heteroskedastic-consistent S.E.
PCRATE	0.0735	0.0332	0.0328
VCRATE	-0.1606	0.0821	0.0828
γ			
Intercept (γ_0)	-2.719	0.824	0.8084
COHORT	-0.4919	0.197	0.1806
BAYAREA	0.7049	0.1422	0.1314
SONOTLA	0.8273	0.1787	0.1729
NORCNTRL	0.542	0.1791	0.1725
BLACK	-0.2554	0.1225	0.1155
AGEREL	0.1754	0.0362	0.0382
DISCHGED	0.2975	0.1832	0.1687
TIMEIN	0.1273	0.0563	0.0549
ROBBER	0.0983	0.0641	0.0572
BURGLAR	-0.0534	0.0371	0.0332
DRUGS	-0.0927	0.043	0.0429
ALCOHOL	-0.1969	0.073	0.0693
DRUGABU	0.112	0.0774	0.0747
SIBCRIM	-0.1221	0.0557	0.0572
NEGLECT	-0.0864	0.0667	0.0646
PCRATE	-0.143	0.0902	0.0826
VCRATE	0.4711	0.2248	0.2133
log-likelihood = -1617.2703; n = 3435			

Table B.2. Negative Binomial Model Results (Second & Third Year Post-Release)

Variable	Parameter Estimate	S.E.	Heteroskedastic-consistent S.E.
β			
β_0	0.8193	0.3477	0.3586
COHORT	-0.1787	0.0458	0.0488
BAYAREA	0.5249	0.0744	0.0749
SONOTLA	0.4934	0.0852	0.0854
NORCNTRL	0.3955	0.0678	0.0677
HISPANIC	0.1597	0.0501	0.0542
BLACK	0.3581	0.049	0.0537
YAGANG	0.0842	0.0301	0.0339
INFRATE	0.0738	0.0122	0.0136
FIRSTADM	-0.1333	0.0481	0.0497
INCAR3PR	0.1492	0.1035	0.1161
BURGLAR	0.0475	0.0119	0.0133
OTHPRO	0.057	0.013	0.0142
MISCHG	0.0456	0.0058	0.0068
DRUGS	0.0573	0.0146	0.0164
AGEFIRST	-0.0164	0.0076	0.0083
ALCOHOL	-0.078	0.0287	0.0282
DRUGABU	0.1105	0.0264	0.0289
ABUSE	-0.0556	0.0327	0.0354
CONTROL	0.0596	0.0238	0.026
JURISFND	-0.2609	0.0804	0.081
FPOVERTY	2.8494	1.2283	1.1837
VCLRATE	-1.0705	0.2811	0.2956

Variable	Parameter Estimate	S.E.	Heteroskedastic-consistent S.E.
γ			
γ_0	-0.6296	0.7437	0.778
BAYAREA	0.3601	0.1598	0.1501
SONOTLA	0.4532	0.1517	0.1425
NORCNTRL	0.2507	0.1187	0.1153
OTHETHN	0.4844	0.2449	0.2738
AGEREL	0.1104	0.0235	0.0259
DISCHGED	0.2562	0.1343	0.1443
FIRSTADM	0.2127	0.1027	0.1009
VIOLENCE	-0.085	0.0362	0.0375
ROBBER	-0.1595	0.0547	0.056
PREVIO	0.1425	0.0537	0.0567
ALCOHOL	-0.0779	0.0566	0.0529
FAMVIO	-0.129	0.0704	0.0762
PARCRIM	0.0952	0.0519	0.0549
JURISFND	-0.2928	0.1751	0.176
FPOVERTY	4.1321	2.7589	2.4777
PCRATE	-0.0808	0.0411	0.0412
LOG-LIKELIHOOD = 1929.0666; N = 3435			

APPENDIX C

MODEL COEFFICIENTS AND GOODNESS OF FIT GRAPHS

Model 1 (Intake Through Month 3)

Model 2 (Month 4 through Month 9)

Model 3 (Month 10 through Month 15)

Model 5 (Month 22 through Month 27)

Model 6 (Month 28 through Month 33)

Model 7 (Month 34 through Month 39)

Model 8 (Month 40 through Month 48)

Data Set: Intake through Month 3

N = 44990 cases.

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke Arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-3.62361	0.2009	-18.03	0.000
	2/6	-5.03113	0.2736	-18.39	0.000
	3/6	-3.53136	0.1343	-26.30	0.000
SEX	1/6	-0.30167	0.1494	-2.02	0.043
	2/6	0.06194	0.1515	0.41	0.683
	3/6	0.02100	0.0744	0.28	0.778
SPLIT	1/6	-0.27835	0.3211	-0.87	0.386
	2/6	-0.01890	0.3387	-0.06	0.956
	3/6	-0.06636	0.1539	-0.43	0.666
PRPRSN	1/6	0.29868	0.0706	4.23	0.000
	2/6	0.01355	0.0986	0.14	0.891
	3/6	0.17694	0.0450	3.93	0.000
ADMITS	1/6	0.12515	0.0580	2.16	0.031
	2/6	0.23232	0.0640	3.63	0.000
	3/6	0.05473	0.0340	1.61	0.107
VIOLENT	1/6	-0.55614	0.1775	-3.13	0.002
	2/6	0.12435	0.1825	0.68	0.496
	3/6	-0.55181	0.0907	-6.08	0.000
DRUG	1/6	0.13337	0.1224	1.09	0.276
	2/6	0.55509	0.1438	3.86	0.000
	3/6	-0.16283	0.0705	-2.31	0.021
OTHER	1/6	-0.64972	0.1967	-3.30	0.001
	2/6	-0.77719	0.2693	-2.89	0.004
	3/6	-1.15355	0.1283	-8.99	0.000
CIRCT2	1/6	1.31924	0.6793	1.94	0.052
	2/6	0.72168	0.5883	1.23	0.220
	3/6	-0.14073	0.1996	-0.71	0.481
CIRCT3	1/6	-0.57332	0.7569	-0.76	0.449
	2/6	0.58807	0.6839	0.86	0.390
	3/6	0.00207	0.3456	0.01	0.995
CIRCT5	1/6	0.89681	0.4775	1.88	0.060
	2/6	0.96726	0.6729	1.44	0.151
	3/6	0.09271	0.2054	0.45	0.652
CIRCT7	1/6	0.47694	0.3705	1.29	0.198
	2/6	-0.15655	0.6155	-0.25	0.799
	3/6	1.03854	0.1990	5.22	0.000
CIRCT8	1/6	-1.73587	1.0333	-1.68	0.093

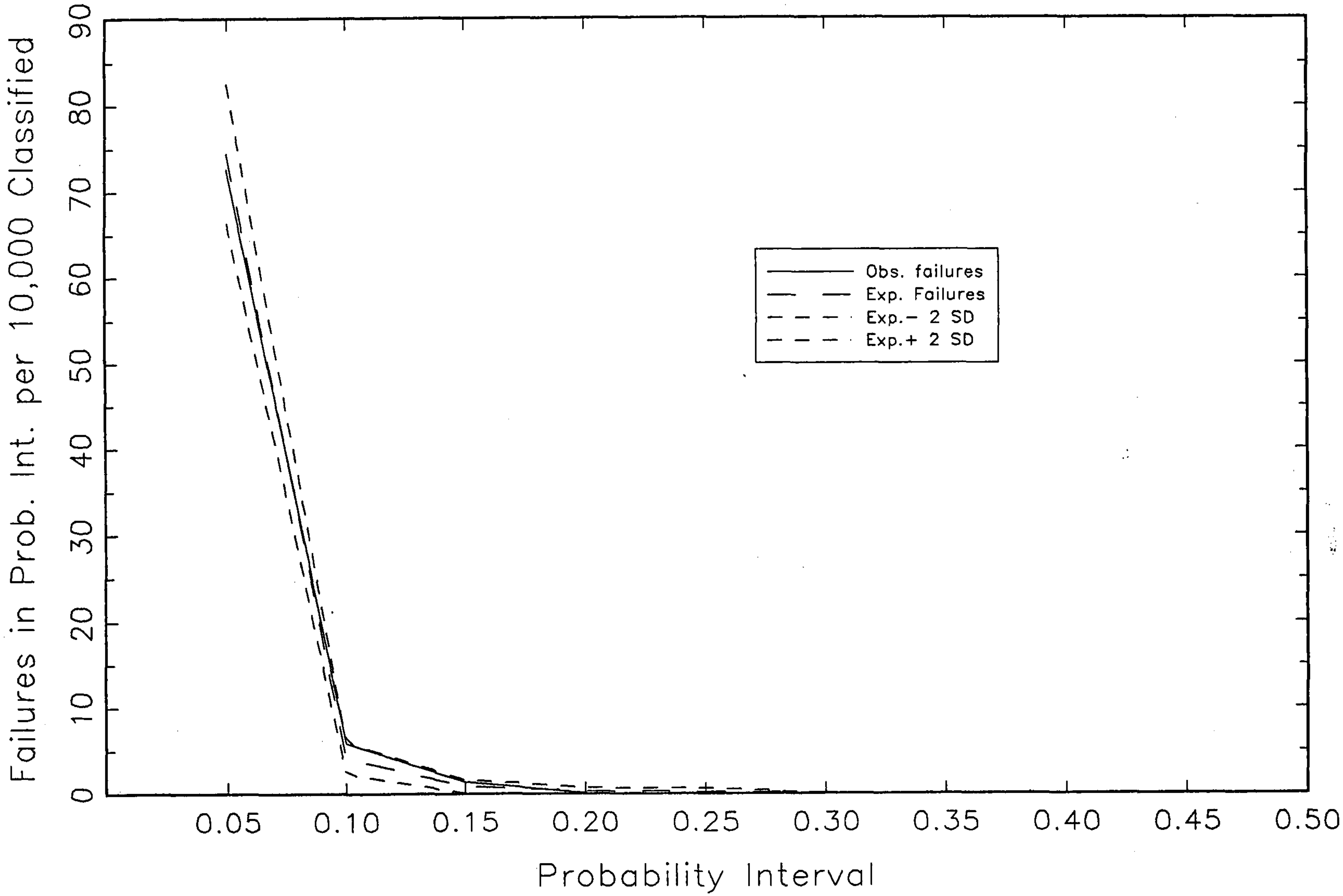
	2/6	-1.08242	1.0632	-1.02	0.309
	3/6	-0.04141	0.3060	-0.14	0.892
CIRCT10	1/6	-0.26873	0.3365	-0.80	0.425
	2/6	-0.02555	0.4283	-0.06	0.952
	3/6	0.43928	0.1695	2.59	0.010
CIRCT11	1/6	-0.55114	0.1873	-2.94	0.003
	2/6	-0.62969	0.2148	-2.93	0.003
	3/6	0.50835	0.1641	3.10	0.002
CIRCT12	1/6	-1.35065	0.4700	-2.87	0.004
	2/6	-0.09963	0.3870	-0.26	0.797
	3/6	0.18686	0.1756	1.06	0.287
CIRCT13	1/6	-0.00833	0.2141	-0.04	0.969
	2/6	0.66349	0.2514	2.64	0.008
	3/6	1.02302	0.1194	8.57	0.000
CIRCT14	1/6	2.43122	0.6356	3.82	0.000
	2/6	1.01038	0.6092	1.66	0.097
	3/6	-0.01004	0.2155	-0.05	0.963
CIRCT15	1/6	-0.10104	0.2366	-0.43	0.669
	2/6	-0.17045	0.2693	-0.63	0.527
	3/6	-0.54263	0.3062	-1.77	0.076
CIRCT16	1/6	-0.06487	0.3993	-0.16	0.871
	2/6	0.09121	0.4031	0.23	0.821
	3/6	0.35844	0.3440	1.04	0.297
CIRCT18	1/6	-0.42148	0.6785	-0.62	0.534
	2/6	1.11324	0.6283	1.77	0.076
	3/6	0.12937	0.1962	0.66	0.510
CIRCT19	1/6	-0.98845	0.4271	-2.31	0.021
	2/6	-0.46807	0.3785	-1.24	0.216
	3/6	1.07804	0.2091	5.15	0.000
CIRCT20	1/6	-1.12031	0.4708	-2.38	0.017
	2/6	-0.44135	0.4892	-0.90	0.367
	3/6	0.14156	0.1962	0.72	0.471
LAGEADM	1/6	-0.20109	0.0587	-3.42	0.001
	2/6	-0.01209	0.0720	-0.17	0.867
	3/6	0.06667	0.0350	1.90	0.057
REGION1	1/6	-2.06351	0.5958	-3.46	0.001
	2/6	-0.87911	0.4895	-1.80	0.072
	3/6	0.33003	0.1595	2.07	0.039
REGION2	1/6	-0.72535	0.2908	-2.49	0.013
	2/6	-0.68541	0.4067	-1.69	0.092
	3/6	-0.32755	0.1876	-1.75	0.081
REGION3	1/6	-1.59223	0.3820	-4.17	0.000
	2/6	-1.64095	0.5375	-3.05	0.002
	3/6	-0.23647	0.1580	-1.50	0.134
REGION4	1/6	0.26095	0.1818	1.44	0.151
	2/6	0.72042	0.2314	3.11	0.002
	3/6	-0.66805	0.1609	-4.15	0.000
LYRSUP2	1/6	-0.85290	0.0717	-11.89	0.000
	2/6	-0.74171	0.0863	-8.59	0.000
	3/6	-0.31213	0.0442	-7.06	0.000
LCOUNTS2	1/6	0.05573	0.1069	0.52	0.602
	2/6	-0.06869	0.1282	-0.54	0.592
	3/6	0.02648	0.0563	0.47	0.638

MEASURES OF FIT:

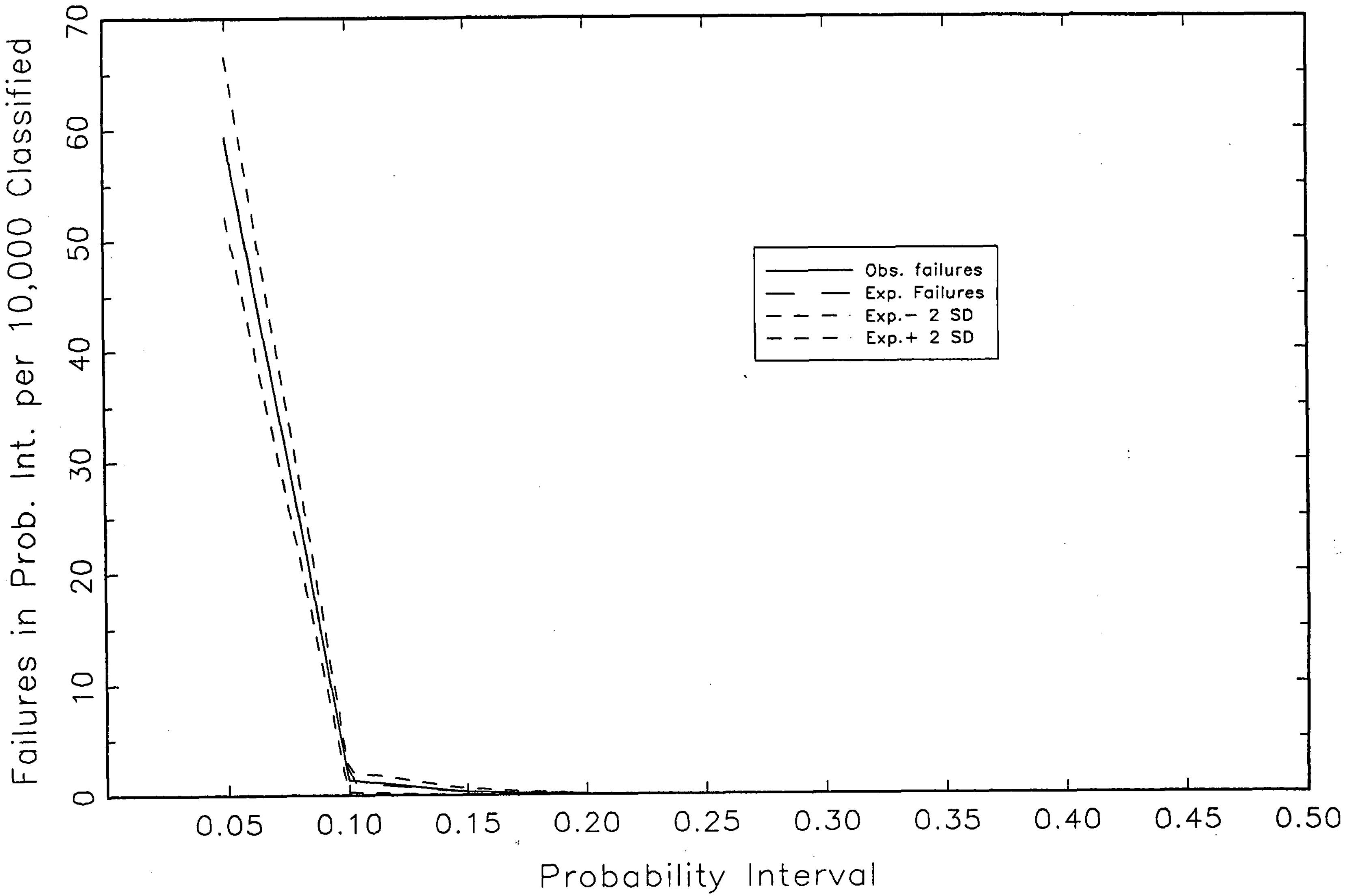
Test	LRX2	df	Prob
Overall	1061.9472	87	0.000
CONSTANT	1313.4654	3	0.000
SEX	4.3615	3	0.225
SPLIT	0.9245	3	0.820
PRPRSN	32.0021	3	0.000
ADMITS	19.6487	3	0.000
VIOLENT	46.9518	3	0.000
DRUG	21.7143	3	0.000
OTHER	98.7298	3	0.000
CIRCT2	5.8035	3	0.122
CIRCT3	1.3230	3	0.724
CIRCT5	5.7438	3	0.125
CIRCT7	28.7467	3	0.000
CIRCT8	3.8650	3	0.276
CIRCT10	7.4440	3	0.059
CIRCT11	27.1665	3	0.000
CIRCT12	9.5330	3	0.023
CIRCT13	79.5273	3	0.000
CIRCT14	17.3203	3	0.001
CIRCT15	3.6717	3	0.299
CIRCT16	1.1629	3	0.762
CIRCT18	3.9547	3	0.266
CIRCT19	34.0553	3	0.000
CIRCT20	7.0327	3	0.071
LAGEADM	15.5911	3	0.001
REGION1	19.6732	3	0.000
REGION2	11.8202	3	0.008
REGION3	28.6240	3	0.000
REGION4	29.4545	3	0.000
LYRSUP2	252.4760	3	0.000
LCOUNTS2	0.7862	3	0.853

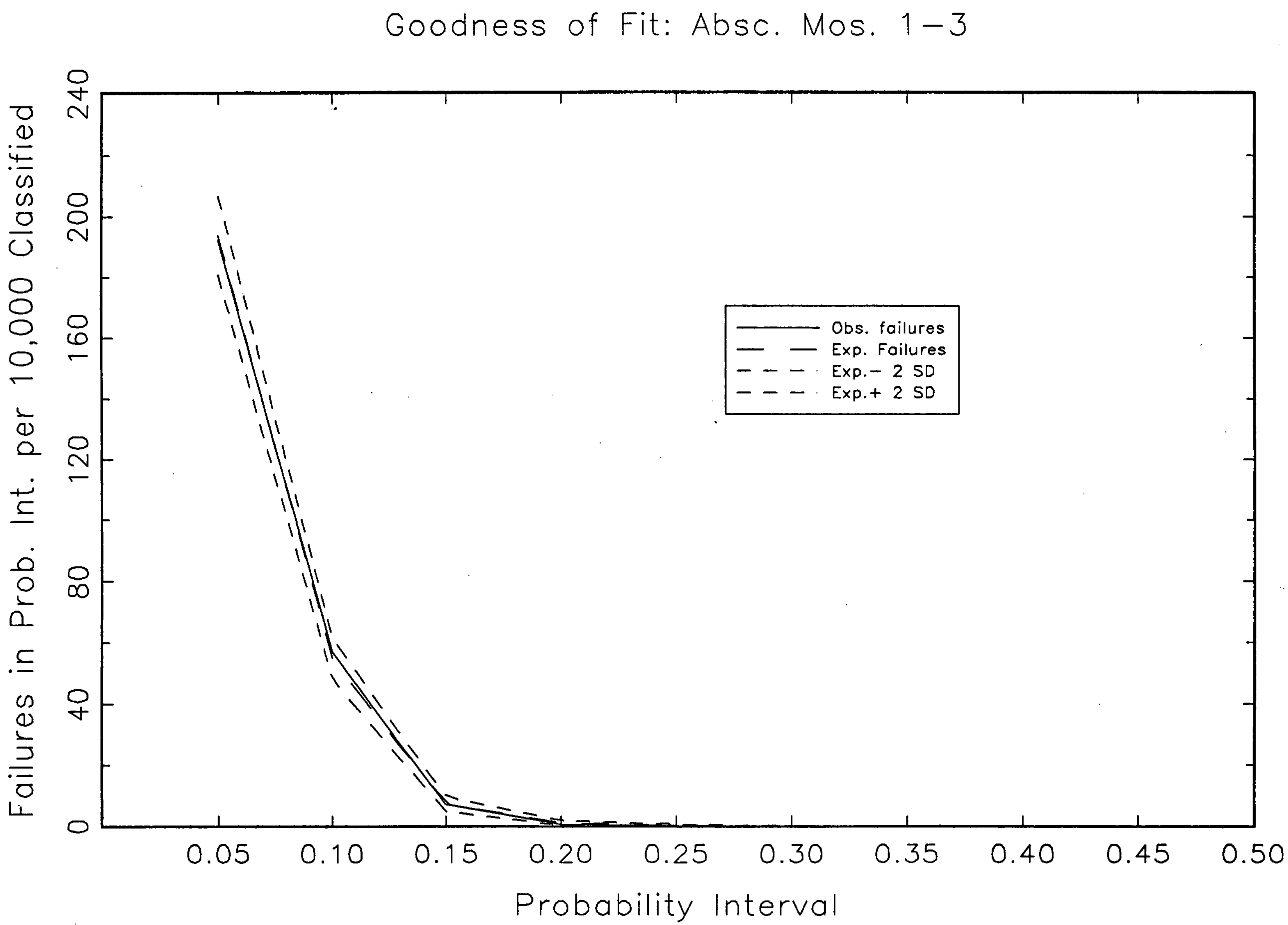
-2 Log Likelihood for full model: 17148.6211
-2 Log likelihood for restricted model: 18210.5683

Goodness of Fit: Rev. Arr. Mos. 1-3



Goodness of Fit: Rev. Tech. Mos. 1-3





Goodness of Fit: Absc. Mos. 1-3
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval

Data Set: Months 4 through 9

N = 50436 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-1.57019	0.0772	-20.33	0.000
	2/6	-2.33244	0.0888	-26.25	0.000
	3/6	-2.02548	0.0798	-25.37	0.000
SEX	1/6	-0.37075	0.0549	-6.75	0.000
	2/6	-0.03337	0.0500	-0.67	0.505
	3/6	-0.28244	0.0493	-5.73	0.000
SPLIT	1/6	0.16336	0.0829	1.97	0.049
	2/6	-0.28097	0.1085	-2.59	0.010
	3/6	0.00848	0.0839	0.10	0.919
PRPRSN	1/6	0.29625	0.0281	10.55	0.000
	2/6	0.16896	0.0328	5.15	0.000
	3/6	0.20682	0.0294	7.04	0.000
ADMITS	1/6	0.23828	0.0213	11.19	0.000
	2/6	0.21773	0.0226	9.63	0.000
	3/6	0.12242	0.0214	5.72	0.000
VIOLENT	1/6	-0.20994	0.0558	-3.76	0.000
	2/6	-0.20989	0.0595	-3.53	0.000
	3/6	-0.49391	0.0528	-9.36	0.000
DRUG	1/6	0.26614	0.0455	5.85	0.000
	2/6	0.40920	0.0463	8.85	0.000
	3/6	-0.12401	0.0443	-2.80	0.005
OTHER	1/6	-0.45683	0.0716	-6.38	0.000
	2/6	-0.56133	0.0774	-7.26	0.000
	3/6	-0.49374	0.0637	-7.76	0.000
CIRCT2	1/6	0.16429	0.1450	1.13	0.257
	2/6	0.41153	0.1963	2.10	0.036
	3/6	-0.14951	0.1219	-1.23	0.220
CIRCT3	1/6	-0.00052	0.2242	-0.00	0.998
	2/6	-0.33950	0.2309	-1.47	0.141
	3/6	-0.03582	0.1748	-0.20	0.838
CIRCT5	1/6	0.05389	0.1221	0.44	0.659
	2/6	-0.12636	0.1372	-0.92	0.357
	3/6	-0.17240	0.1216	-1.42	0.156
CIRCT7	1/6	0.47033	0.1543	3.05	0.002
	2/6	0.01807	0.1466	0.12	0.902
	3/6	0.66748	0.1132	5.90	0.000
CIRCT8	1/6	0.40123	0.1734	2.31	0.021
	2/6	-0.49797	0.1972	-2.53	0.012
	3/6	-0.42297	0.1732	-2.44	0.015

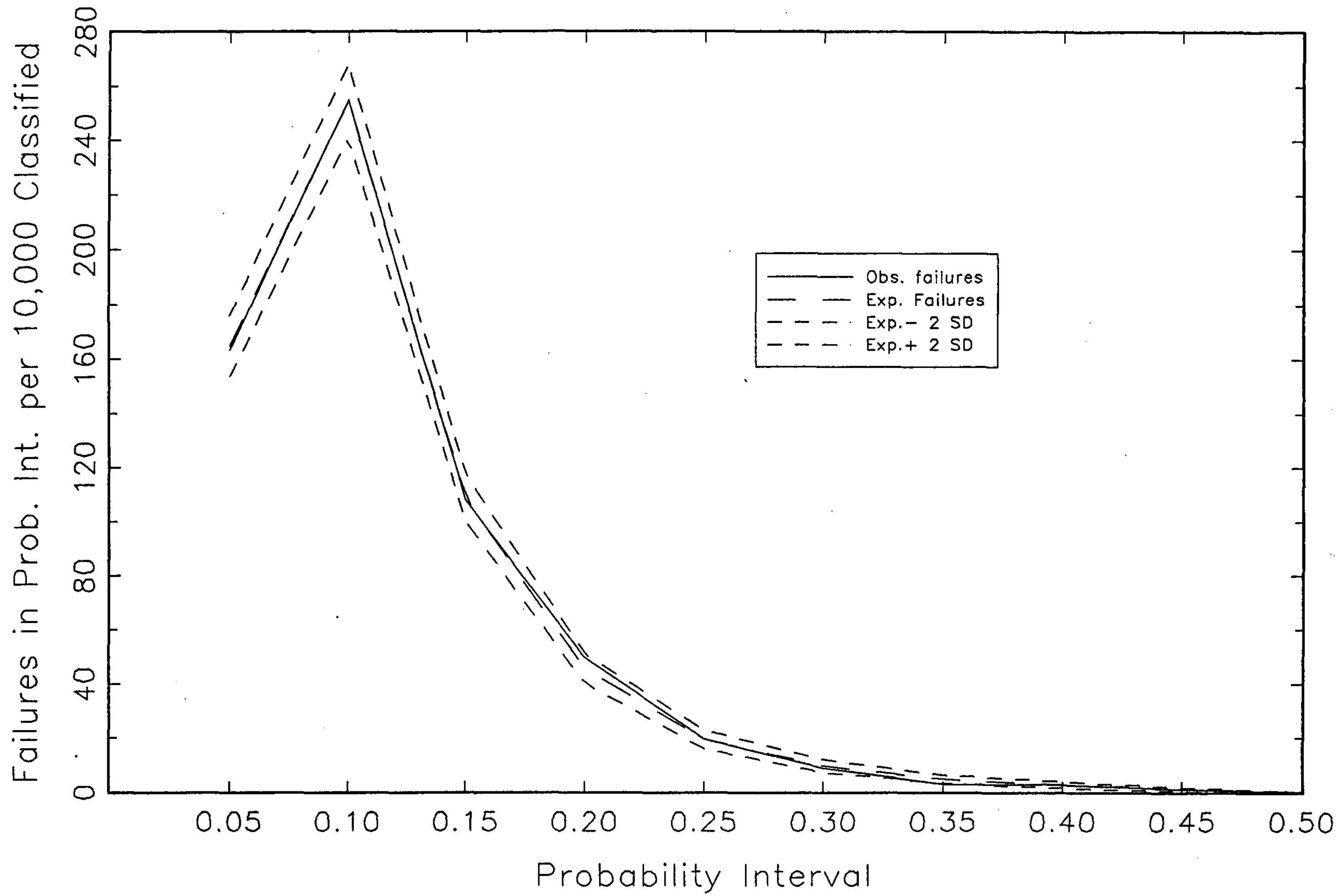
CIRCT10	1/6	-0.35768	0.1159	-3.09	0.002
	2/6	0.47538	0.1078	4.41	0.000
	3/6	0.17730	0.1020	1.74	0.082
CIRCT11	1/6	-0.26875	0.0759	-3.54	0.000
	2/6	-0.66535	0.0800	-8.32	0.000
	3/6	1.00503	0.1061	9.48	0.000
CIRCT12	1/6	-0.41485	0.1136	-3.65	0.000
	2/6	-0.27050	0.1280	-2.11	0.035
	3/6	0.27810	0.0975	2.85	0.004
CIRCT13	1/6	-0.26429	0.0876	-3.02	0.003
	2/6	0.76154	0.0829	9.18	0.000
	3/6	0.73494	0.0753	9.77	0.000
CIRCT14	1/6	0.48597	0.1476	3.29	0.001
	2/6	0.74208	0.2005	3.70	0.000
	3/6	0.24741	0.1227	2.02	0.044
CIRCT15	1/6	0.04465	0.0963	0.46	0.643
	2/6	-0.45822	0.1093	-4.19	0.000
	3/6	0.17573	0.1614	1.09	0.276
CIRCT16	1/6	0.03350	0.1705	0.20	0.844
	2/6	0.27298	0.1466	1.86	0.063
	3/6	1.59863	0.1612	9.92	0.000
CIRCT18	1/6	-0.11612	0.1213	-0.96	0.338
	2/6	-0.25886	0.1333	-1.94	0.052
	3/6	-0.18372	0.1155	-1.59	0.112
CIRCT19	1/6	-0.13432	0.1238	-1.08	0.278
	2/6	-0.22750	0.1250	-1.82	0.069
	3/6	1.46557	0.1301	11.27	0.000
CIRCT20	1/6	-0.24330	0.1223	-1.99	0.047
	2/6	-0.09101	0.1375	-0.66	0.508
	3/6	0.04545	0.1183	0.38	0.701
LAGEADM	1/6	-0.36832	0.0206	-17.84	0.000
	2/6	-0.16419	0.0222	-7.39	0.000
	3/6	-0.08715	0.0202	-4.31	0.000
REGION1	1/6	-0.63738	0.1099	-5.80	0.000
	2/6	-0.99067	0.1506	-6.58	0.000
	3/6	0.10198	0.0948	1.08	0.282
REGION2	1/6	-0.90118	0.1244	-7.24	0.000
	2/6	-0.20560	0.1132	-1.82	0.069
	3/6	-0.12369	0.1041	-1.19	0.235
REGION3	1/6	-0.46496	0.0926	-5.02	0.000
	2/6	-0.18449	0.1009	-1.83	0.068
	3/6	-0.16420	0.0898	-1.83	0.068
REGION4	1/6	-0.11274	0.0738	-1.53	0.127
	2/6	0.36945	0.0792	4.66	0.000
	3/6	-1.18403	0.1062	-11.15	0.000
LYRSUP2	1/6	-0.38192	0.0377	-10.13	0.000
	2/6	-0.45236	0.0404	-11.19	0.000
	3/6	-0.25389	0.0338	-7.52	0.000
LCOUNTS2	1/6	0.24478	0.0346	7.08	0.000
	2/6	0.09238	0.0389	2.37	0.018
	3/6	0.11358	0.0333	3.41	0.001
CATSUP	1/6	0.59118	0.0771	7.67	0.000
	2/6	0.64102	0.0786	8.16	0.000
	3/6	-0.00492	0.0833	-0.06	0.953

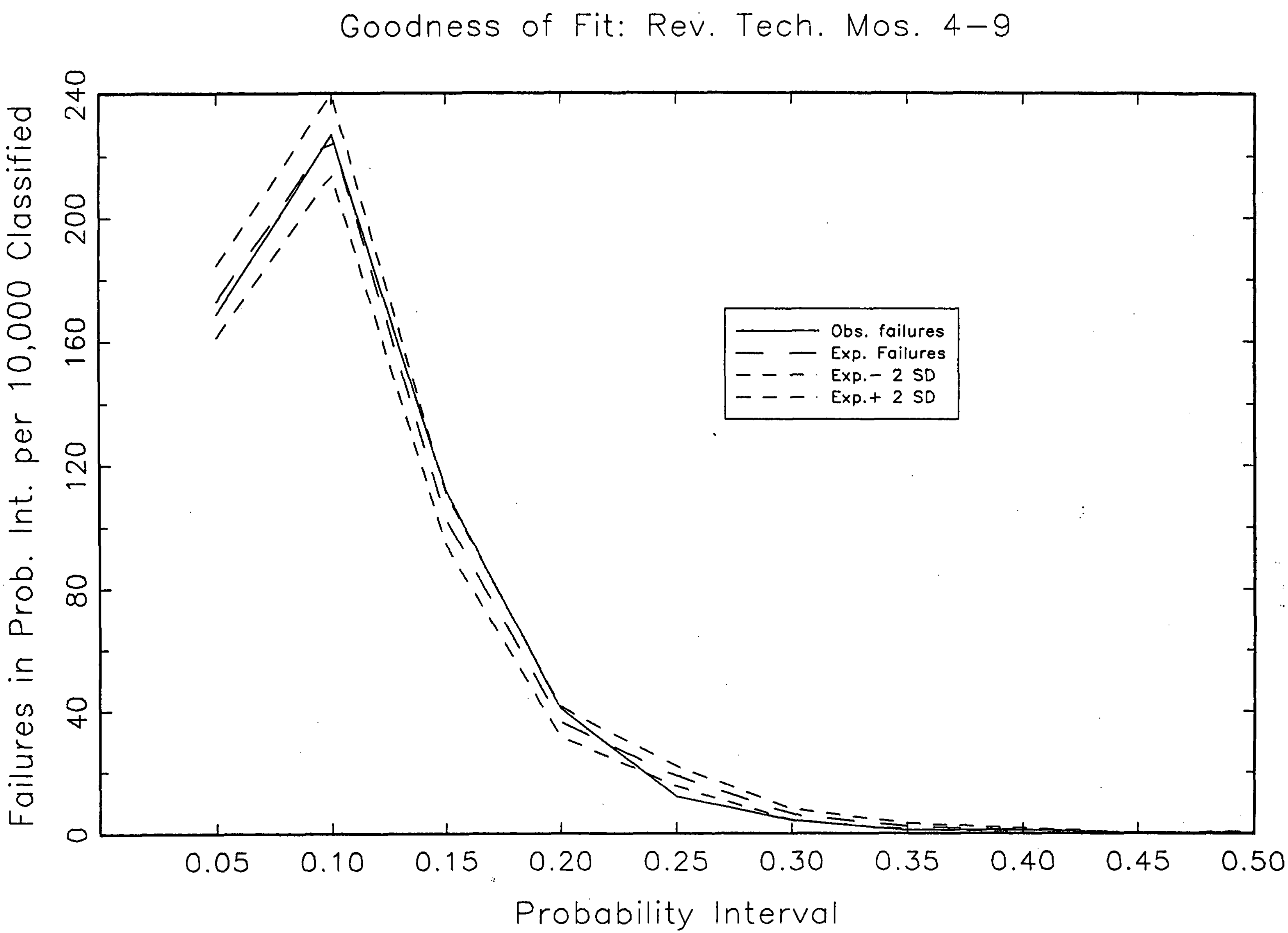
MEASURES OF FIT:

Test	LRX2	df	Prob
Overall	3828.7179	90	0.000
CONSTANT	1484.7690	3	0.000
SEX	73.6169	3	0.000
SPLIT	11.7840	3	0.008
PRPRSN	146.8212	3	0.000
ADMITS	203.2815	3	0.000
VIOLENT	104.5414	3	0.000
DRUG	119.4354	3	0.000
OTHER	135.5811	3	0.000
CIRCT2	7.4810	3	0.058
CIRCT3	2.1839	3	0.535
CIRCT5	3.0759	3	0.380
CIRCT7	41.7391	3	0.000
CIRCT8	18.6783	3	0.000
CIRCT10	34.3448	3	0.000
CIRCT11	180.1687	3	0.000
CIRCT12	27.9473	3	0.000
CIRCT13	186.1587	3	0.000
CIRCT14	25.1969	3	0.000
CIRCT15	19.8706	3	0.000
CIRCT16	99.6733	3	0.000
CIRCT18	6.4967	3	0.090
CIRCT19	137.4241	3	0.000
CIRCT20	4.5522	3	0.208
LAGEADM	357.9331	3	0.000
REGION1	76.2065	3	0.000
REGION2	54.3718	3	0.000
REGION3	29.0283	3	0.000
REGION4	155.0205	3	0.000
LYRSUP2	247.5847	3	0.000
LCOUNTS2	60.0751	3	0.000
CATSUP	114.1689	3	0.000

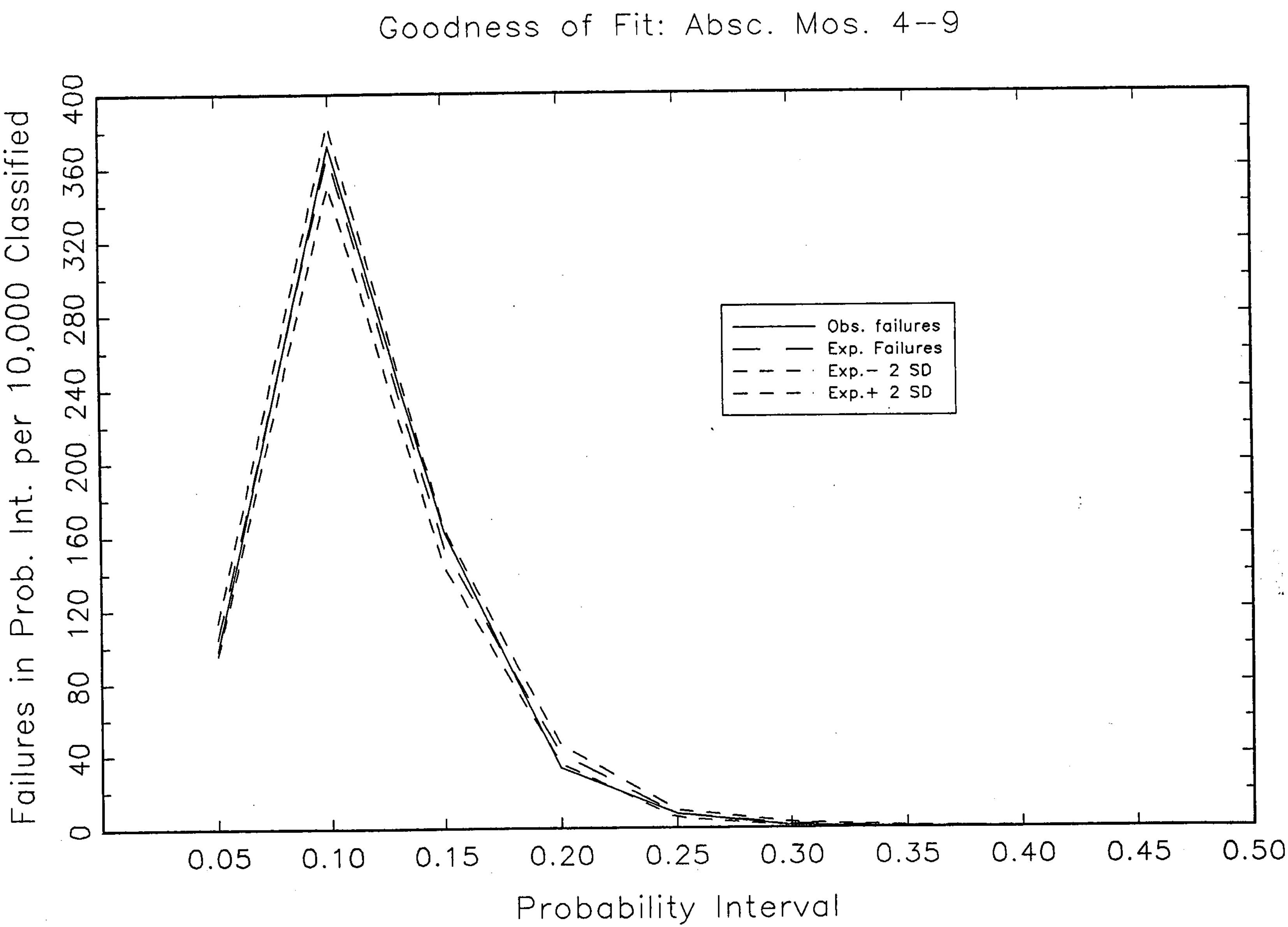
-2 Log Likelihood for full model: 64964.2815
-2 Log likelihood for restricted model: 68792.9994

Goodness of Fit: Rev. Arr. Mos. 4–9





Goodness of Fit: Rev. Tech. Mos. 4–9
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.– 2 SD
Exp.+ 2 SD


Goodness of Fit: Absc. Mos. 4--9
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
280
320
360
400
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD

Data Set: Months 10 through 15

N = 38366 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-1.29178	0.0975	-13.25	0.000
	2/6	-1.94255	0.1041	-18.66	0.000
	3/6	-2.27495	0.1090	-20.87	0.000
SEX	1/6	-0.54855	0.0681	-8.06	0.000
	2/6	-0.24117	0.0572	-4.22	0.000
	3/6	-0.23468	0.0608	-3.86	0.000
SPLIT	1/6	0.42525	0.0850	5.00	0.000
	2/6	0.05059	0.0985	0.51	0.608
	3/6	0.29163	0.0910	3.21	0.001
PRPRSN	1/6	0.39288	0.0313	12.57	0.000
	2/6	0.18356	0.0368	4.99	0.000
	3/6	0.16860	0.0382	4.41	0.000
ADMITS	1/6	0.19332	0.0195	9.90	0.000
	2/6	0.16720	0.0202	8.28	0.000
	3/6	0.14491	0.0209	6.93	0.000
VIOLENT	1/6	-0.08207	0.0613	-1.34	0.181
	2/6	-0.05897	0.0596	-0.99	0.322
	3/6	-0.40215	0.0634	-6.34	0.000
DRUG	1/6	0.14605	0.0549	2.66	0.008
	2/6	0.23468	0.0521	4.51	0.000
	3/6	-0.18497	0.0580	-3.19	0.001
OTHER	1/6	-0.18290	0.0818	-2.24	0.025
	2/6	-0.38109	0.0838	-4.55	0.000
	3/6	-0.25606	0.0777	-3.30	0.001
CIRCT2	1/6	-0.14969	0.1646	-0.91	0.363
	2/6	0.10102	0.1716	0.59	0.556
	3/6	-0.11112	0.1263	-0.88	0.379
CIRCT3	1/6	-0.15224	0.2199	-0.69	0.489
	2/6	-0.31330	0.1997	-1.57	0.117
	3/6	0.56890	0.2407	2.36	0.018
CIRCT5	1/6	-0.03729	0.1310	-0.28	0.776
	2/6	0.10829	0.1202	0.90	0.368
	3/6	-0.27017	0.1352	-2.00	0.046
CIRCT7	1/6	0.09819	0.1664	0.59	0.555
	2/6	-0.83652	0.1728	-4.84	0.000
	3/6	1.21642	0.1776	6.85	0.000
CIRCT8	1/6	0.17038	0.1837	0.93	0.354
	2/6	-0.44479	0.1760	-2.53	0.011
	3/6	0.32932	0.2334	1.41	0.158

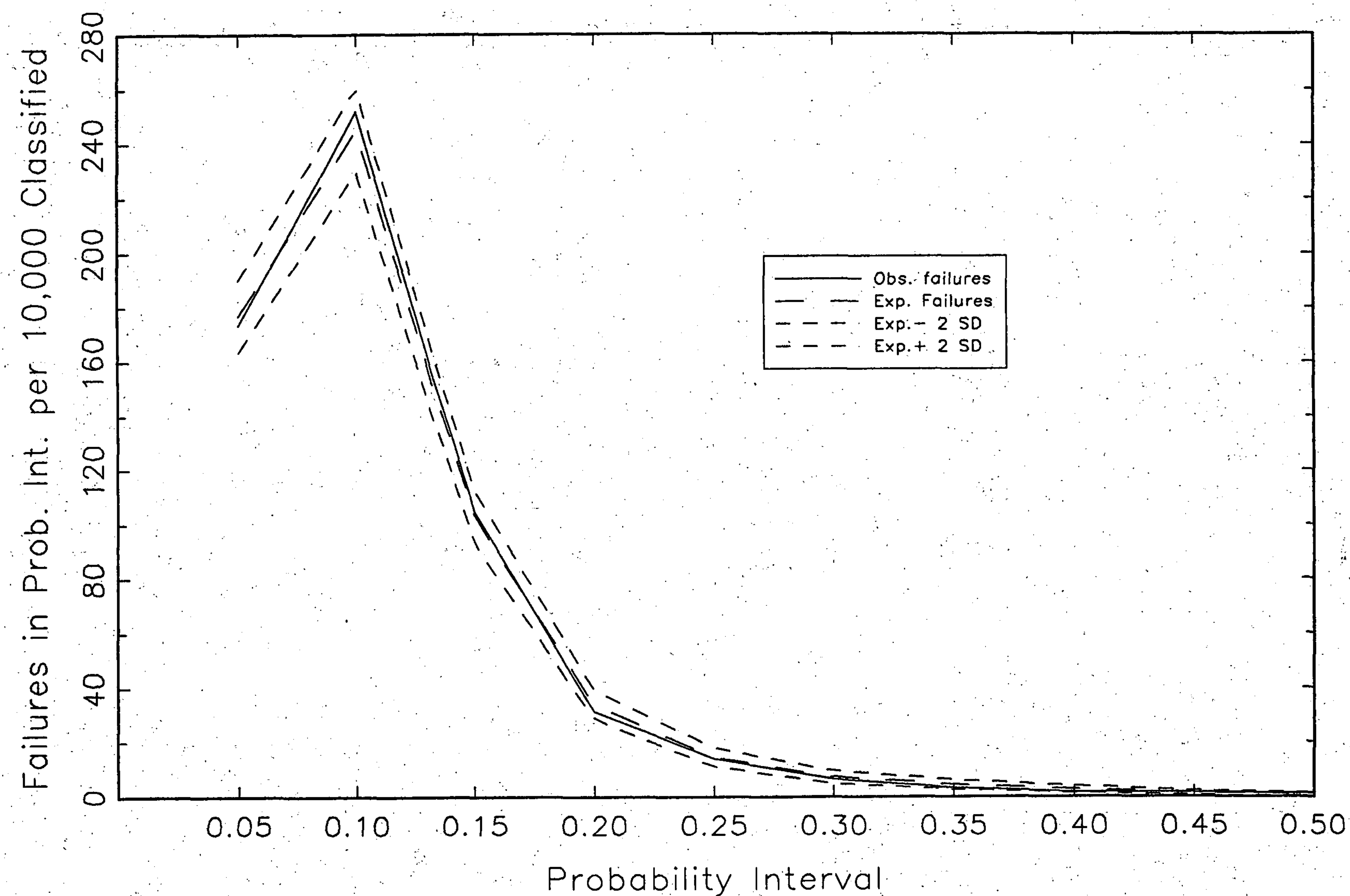
CIRCT10	1/6	-0.39356	0.1350	-2.92	0.004
	2/6	0.51587	0.1148	4.49	0.000
	3/6	0.48607	0.1174	4.14	0.000
CIRCT11	1/6	-0.06344	0.0966	-0.66	0.511
	2/6	-0.43216	0.0972	-4.44	0.000
	3/6	1.10100	0.1483	7.42	0.000
CIRCT12	1/6	-0.46565	0.1360	-3.42	0.001
	2/6	-0.21201	0.1353	-1.57	0.117
	3/6	-0.00903	0.1339	-0.07	0.946
CIRCT13	1/6	-0.51112	0.1126	-4.54	0.000
	2/6	0.26679	0.1020	2.62	0.009
	3/6	0.33056	0.1041	3.18	0.001
CIRCT14	1/6	0.16003	0.1746	0.92	0.359
	2/6	0.41094	0.1851	2.22	0.026
	3/6	-0.00114	0.1437	-0.01	0.994
CIRCT15	1/6	0.17925	0.1100	1.63	0.103
	2/6	-0.09510	0.1121	-0.85	0.396
	3/6	0.13945	0.2227	0.63	0.531
CIRCT16	1/6	-0.13246	0.2398	-0.55	0.581
	2/6	0.50306	0.1681	2.99	0.003
	3/6	2.02291	0.2049	9.87	0.000
CIRCT18	1/6	0.03285	0.1201	0.27	0.784
	2/6	-0.46002	0.1304	-3.53	0.000
	3/6	-0.28645	0.1270	-2.26	0.024
CIRCT19	1/6	-0.00134	0.1405	-0.01	0.992
	2/6	-0.05527	0.1366	-0.40	0.686
	3/6	1.74267	0.1667	10.45	0.000
CIRCT20	1/6	-0.37141	0.1471	-2.53	0.012
	2/6	-0.15725	0.1509	-1.04	0.298
	3/6	-0.12650	0.1542	-0.82	0.412
LAGEADM	1/6	-0.44387	0.0238	-18.66	0.000
	2/6	-0.18506	0.0239	-7.75	0.000
	3/6	-0.13451	0.0254	-5.29	0.000
REGION1	1/6	-0.42440	0.1241	-3.42	0.001
	2/6	-0.42515	0.1355	-3.14	0.002
	3/6	0.42941	0.1099	3.91	0.000
REGION2	1/6	-0.39128	0.1342	-2.92	0.004
	2/6	0.27578	0.1145	2.41	0.016
	3/6	-0.76500	0.1678	-4.56	0.000
REGION3	1/6	-0.20718	0.1019	-2.03	0.042
	2/6	0.14586	0.1015	1.44	0.151
	3/6	0.16522	0.1052	1.57	0.116
REGION4	1/6	-0.23823	0.0901	-2.64	0.008
	2/6	0.08246	0.0914	0.90	0.367
	3/6	-1.41099	0.1452	-9.72	0.000
LYRSUP2	1/6	-0.38386	0.0536	-7.16	0.000
	2/6	-0.49059	0.0552	-8.88	0.000
	3/6	-0.18928	0.0543	-3.48	0.000
LCOUNTS2	1/6	0.14731	0.0401	3.67	0.000
	2/6	0.07327	0.0415	1.76	0.078
	3/6	0.14247	0.0408	3.50	0.000
CATSUP	1/6	-0.21517	0.0758	-2.84	0.005
	2/6	-0.03254	0.0725	-0.45	0.653
	3/6	0.29376	0.0775	3.79	0.000

MEASURES OF FIT:

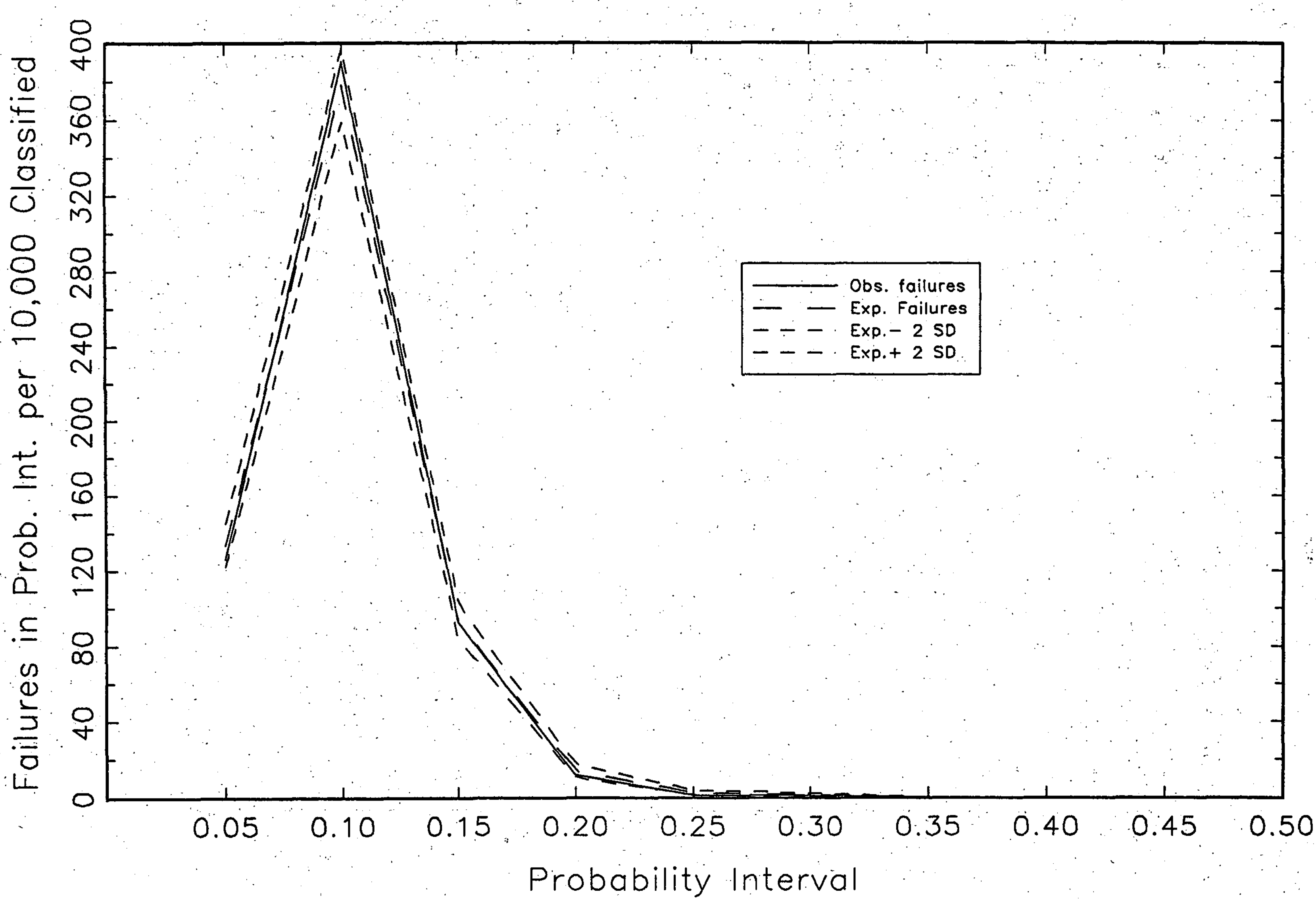
Test	LRX2	df	Prob
Overall	2242.0239	90	0.000
CONSTANT	833.6324	3	0.000
SEX	89.2726	3	0.000
SPLIT	32.0315	3	0.000
PRPRSN	169.6527	3	0.000
ADMITS	165.4276	3	0.000
VIOLENT	41.2485	3	0.000
DRUG	39.2726	3	0.000
OTHER	32.7923	3	0.000
CIRCT2	1.9779	3	0.577
CIRCT3	9.1465	3	0.027
CIRCT5	5.1670	3	0.160
CIRCT7	74.6326	3	0.000
CIRCT8	9.9822	3	0.019
CIRCT10	47.2831	3	0.000
CIRCT11	78.6360	3	0.000
CIRCT12	13.4992	3	0.004
CIRCT13	40.8528	3	0.000
CIRCT14	5.5626	3	0.135
CIRCT15	3.9989	3	0.262
CIRCT16	103.6969	3	0.000
CIRCT18	17.0245	3	0.001
CIRCT19	110.8994	3	0.000
CIRCT20	7.4411	3	0.059
LAGEADM	398.9682	3	0.000
REGION1	39.6582	3	0.000
REGION2	36.5006	3	0.000
REGION3	9.4077	3	0.024
REGION4	101.4108	3	0.000
LYRSUP2	127.0747	3	0.000
LCOUNTS2	25.1397	3	0.000
CATSUP	24.3993	3	0.000

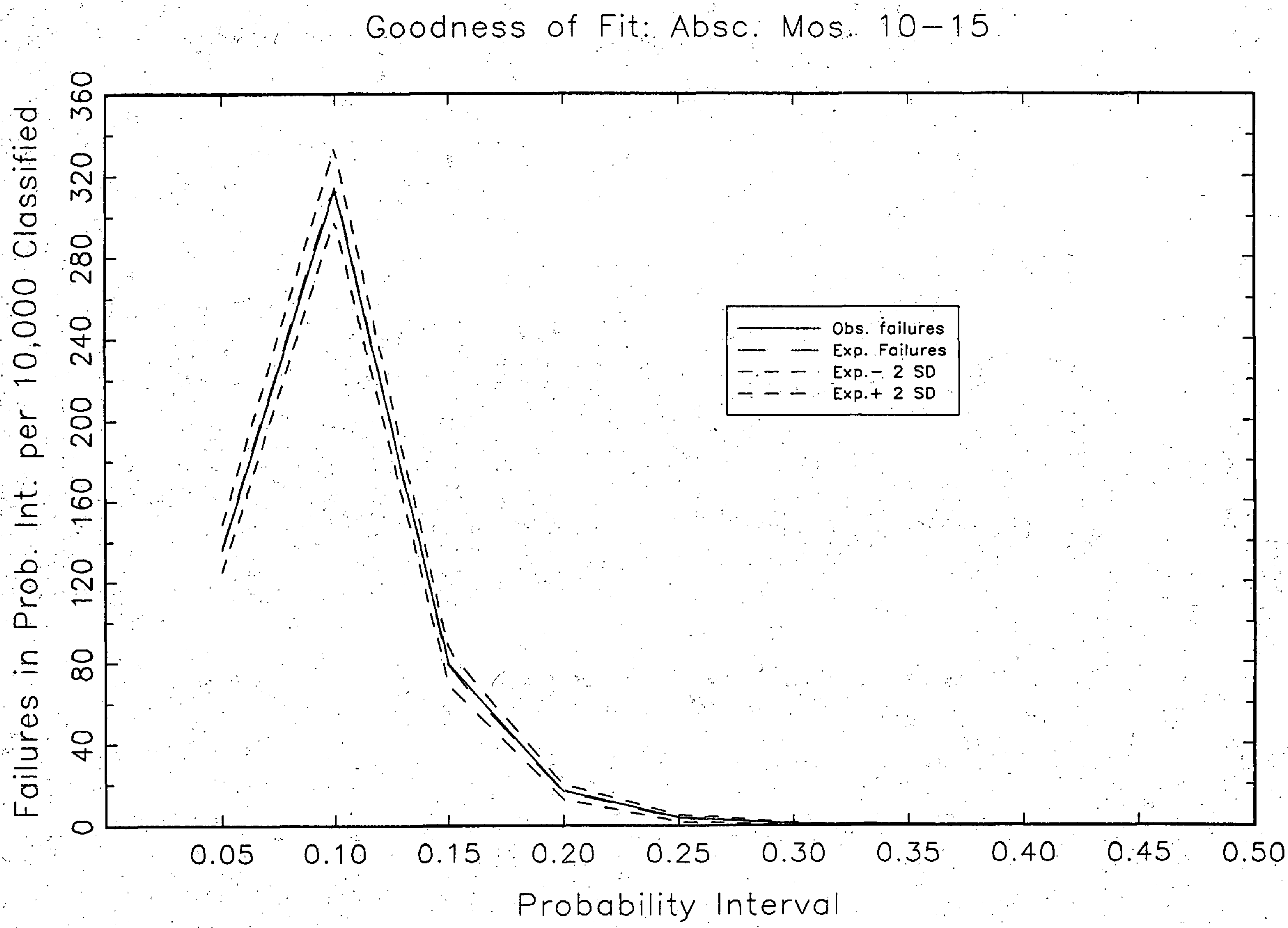
-2 Log Likelihood for full model: 48308.3560
-2 Log likelihood for restricted model: 50550.3799

Goodness of Fit: Rev. Arr. Mos. 10-15



Goodness of Fit: Rev. Tech. Mos. 10–15




Goodness of Fit: Absc. Mos. 10-15
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
280
320
360
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD

Data Set: Months 22 through 27

N = 36475 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

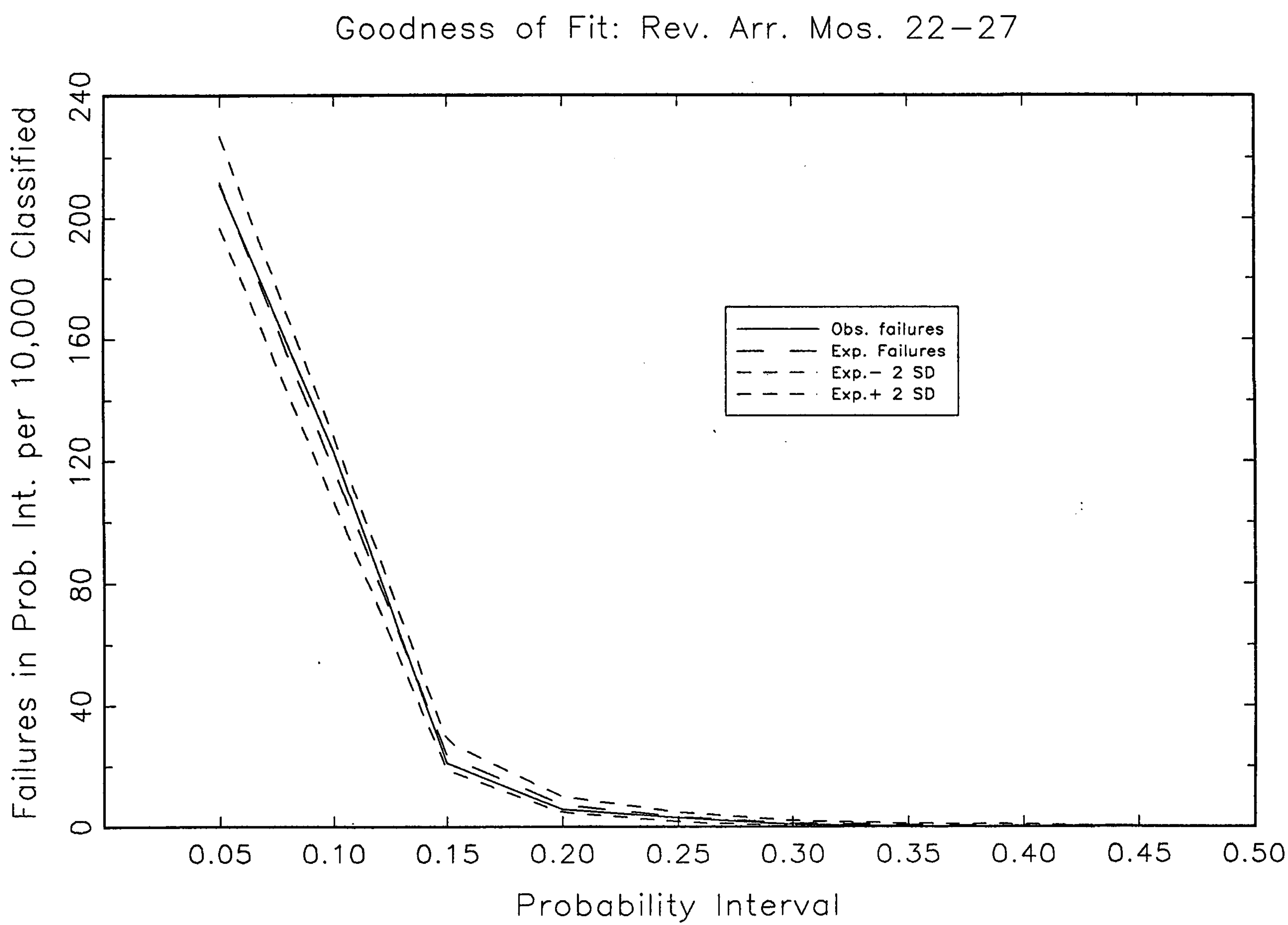
Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-1.63052	0.1516	-10.75	0.000
	2/6	-2.45617	0.1643	-14.95	0.000
	3/6	-2.31956	0.1497	-15.49	0.000
SEX	1/6	-0.62192	0.0893	-6.96	0.000
	2/6	-0.22429	0.0704	-3.18	0.001
	3/6	-0.26468	0.0657	-4.03	0.000
SPLIT	1/6	0.42352	0.0970	4.36	0.000
	2/6	0.04717	0.1123	0.42	0.674
	3/6	0.30879	0.0925	3.34	0.001
PRPRSN	1/6	0.34054	0.0413	8.25	0.000
	2/6	0.18918	0.0477	3.96	0.000
	3/6	0.16811	0.0459	3.66	0.000
ADMITS	1/6	0.18751	0.0216	8.69	0.000
	2/6	0.16743	0.0218	7.67	0.000
	3/6	0.11076	0.0210	5.27	0.000
VIOLENT	1/6	0.12014	0.0775	1.55	0.121
	2/6	-0.05733	0.0759	-0.76	0.450
	3/6	-0.44792	0.0711	-6.30	0.000
DRUG	1/6	0.15063	0.0713	2.11	0.035
	2/6	0.15079	0.0640	2.36	0.018
	3/6	-0.38707	0.0638	-6.07	0.000
OTHER	1/6	-0.00399	0.1071	-0.04	0.970
	2/6	-0.60622	0.1234	-4.91	0.000
	3/6	-0.43607	0.0940	-4.64	0.000
CIRCT2	1/6	-0.89421	0.2271	-3.94	0.000
	2/6	-0.06639	0.2128	-0.31	0.755
	3/6	-0.11104	0.1288	-0.86	0.389
CIRCT3	1/6	0.51929	0.2833	1.83	0.067
	2/6	-0.35946	0.2328	-1.54	0.123
	3/6	0.76584	0.2762	2.77	0.006
CIRCT5	1/6	-0.29468	0.1468	-2.01	0.045
	2/6	-0.09233	0.1275	-0.72	0.469
	3/6	-0.14693	0.1316	-1.12	0.264
CIRCT7	1/6	0.47204	0.2585	1.83	0.068
	2/6	-0.70755	0.2083	-3.40	0.001
	3/6	1.60602	0.2226	7.22	0.000
CIRCT8	1/6	0.85004	0.2726	3.12	0.002
	2/6	-0.75966	0.2622	-2.90	0.004
	3/6	0.37645	0.3036	1.24	0.215

CIRCT10	1/6	-0.81515	0.1834	-4.45	0.000
	2/6	0.64742	0.1462	4.43	0.000
	3/6	0.27526	0.1495	1.84	0.066
CIRCT11	1/6	0.21435	0.1324	1.62	0.105
	2/6	-0.07371	0.1314	-0.56	0.575
	3/6	0.84469	0.1649	5.12	0.000
CIRCT12	1/6	-0.43336	0.1957	-2.21	0.027
	2/6	0.15913	0.1907	0.83	0.404
	3/6	0.35351	0.1697	2.08	0.037
CIRCT13	1/6	-0.76840	0.1390	-5.53	0.000
	2/6	0.12592	0.1353	0.93	0.352
	3/6	0.38318	0.1200	3.19	0.001
CIRCT14	1/6	-0.25758	0.2186	-1.18	0.239
	2/6	0.16836	0.2435	0.69	0.489
	3/6	-0.12110	0.1541	-0.79	0.432
CIRCT15	1/6	0.38457	0.1542	2.49	0.013
	2/6	0.28154	0.1469	1.92	0.055
	3/6	-0.02155	0.2541	-0.08	0.932
CIRCT16	1/6	0.56738	0.2675	2.12	0.034
	2/6	0.10997	0.2966	0.37	0.711
	3/6	1.52937	0.2608	5.86	0.000
CIRCT18	1/6	-0.23388	0.1439	-1.63	0.104
	2/6	-0.64135	0.1472	-4.36	0.000
	3/6	-0.13613	0.1295	-1.05	0.293
CIRCT19	1/6	-0.01801	0.1824	-0.10	0.921
	2/6	-0.16265	0.1795	-0.91	0.365
	3/6	1.69773	0.1629	10.42	0.000
CIRCT20	1/6	-0.57801	0.1701	-3.40	0.001
	2/6	0.42909	0.1542	2.78	0.005
	3/6	0.41718	0.1440	2.90	0.004
LAGEADM	1/6	-0.37992	0.0307	-12.39	0.000
	2/6	-0.18915	0.0301	-6.28	0.000
	3/6	-0.19557	0.0273	-7.16	0.000
REGION1	1/6	-0.20970	0.1460	-1.44	0.151
	2/6	-0.12791	0.1783	-0.72	0.473
	3/6	0.86892	0.1264	6.88	0.000
REGION2	1/6	-1.01447	0.2227	-4.56	0.000
	2/6	0.46083	0.1518	3.04	0.002
	3/6	-0.75196	0.2193	-3.43	0.001
REGION3	1/6	-0.17463	0.1157	-1.51	0.131
	2/6	0.50301	0.1207	4.17	0.000
	3/6	0.30833	0.1186	2.60	0.009
REGION4	1/6	-0.51681	0.1170	-4.42	0.000
	2/6	0.03597	0.1227	0.29	0.769
	3/6	-0.97608	0.1541	-6.33	0.000
LYRSUP2	1/6	-0.49442	0.0841	-5.88	0.000
	2/6	-0.58071	0.0848	-6.85	0.000
	3/6	-0.36661	0.0773	-4.74	0.000
LCOUNTS2	1/6	0.12533	0.0490	2.56	0.011
	2/6	-0.01654	0.0516	-0.32	0.749
	3/6	0.07217	0.0441	1.64	0.102
CATSUP	1/6	-0.04687	0.0857	-0.55	0.584
	2/6	0.42213	0.0831	5.08	0.000
	3/6	0.41099	0.0766	5.37	0.000

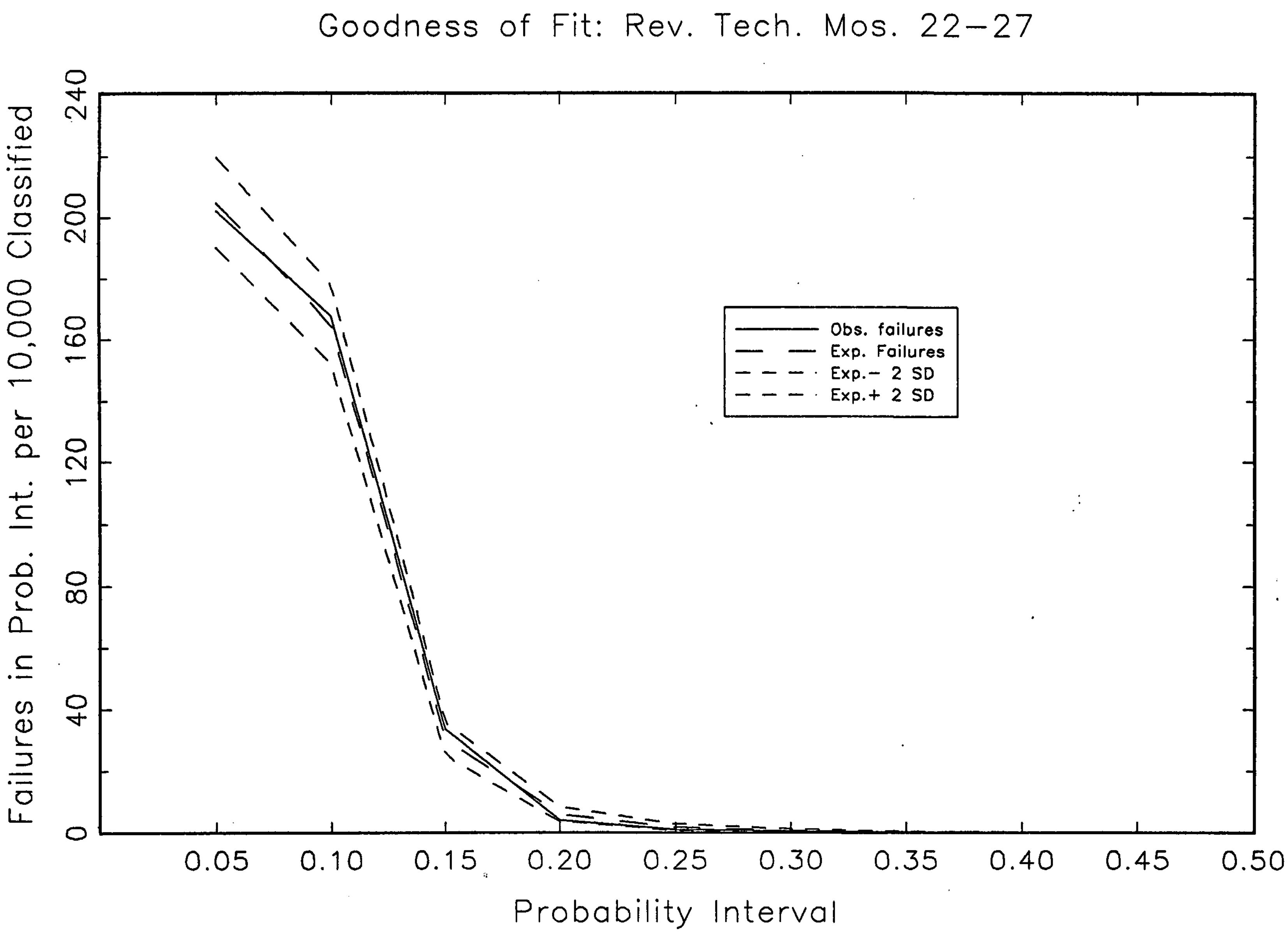
MEASURES OF FIT:

Test	LRX2	df	Prob
Overall	1906.0884	90	0.000
CONSTANT	526.7718	3	0.000
SEX	70.4044	3	0.000
SPLIT	28.1152	3	0.000
PRPRSN	83.9402	3	0.000
ADMITS	134.8817	3	0.000
VIOLENT	43.4816	3	0.000
DRUG	49.4271	3	0.000
OTHER	43.8988	3	0.000
CIRCT2	15.8809	3	0.001
CIRCT3	13.7835	3	0.003
CIRCT5	5.3375	3	0.149
CIRCT7	68.8828	3	0.000
CIRCT8	20.5039	3	0.000
CIRCT10	44.5252	3	0.000
CIRCT11	28.9510	3	0.000
CIRCT12	10.5345	3	0.015
CIRCT13	43.8178	3	0.000
CIRCT14	2.5093	3	0.474
CIRCT15	9.3972	3	0.024
CIRCT16	37.3543	3	0.000
CIRCT18	21.4925	3	0.000
CIRCT19	110.8754	3	0.000
CIRCT20	28.8646	3	0.000
LAGEADM	223.3054	3	0.000
REGION1	52.0406	3	0.000
REGION2	43.2577	3	0.000
REGION3	26.5461	3	0.000
REGION4	57.9566	3	0.000
LYRSUP2	95.0791	3	0.000
LCOUNTS2	9.0415	3	0.029
CATSUP	53.0128	3	0.000

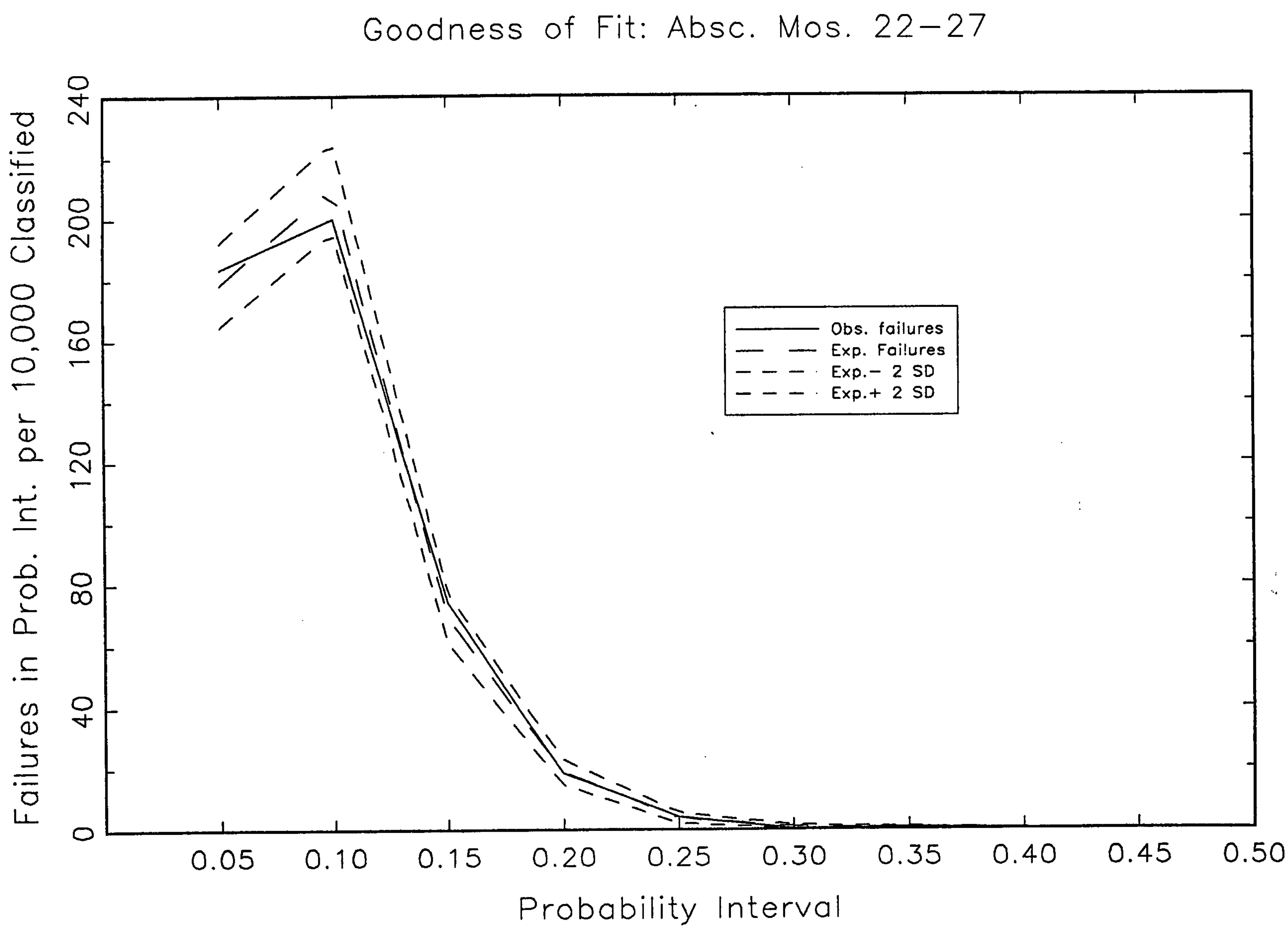
-2 Log Likelihood for full model: 35590.0220
-2 Log likelihood for restricted model: 37496.1104



Goodness of Fit: Rev. Arr. Mos. 22-27
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval



Goodness of Fit: Rev. Tech. Mos. 22–27
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.– 2 SD
Exp.+ 2 SD


Goodness of Fit: Absc. Mos. 22−27
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.− 2 SD
Exp.+ 2 SD

Data Set: Months 28 through 33

N = 17525 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

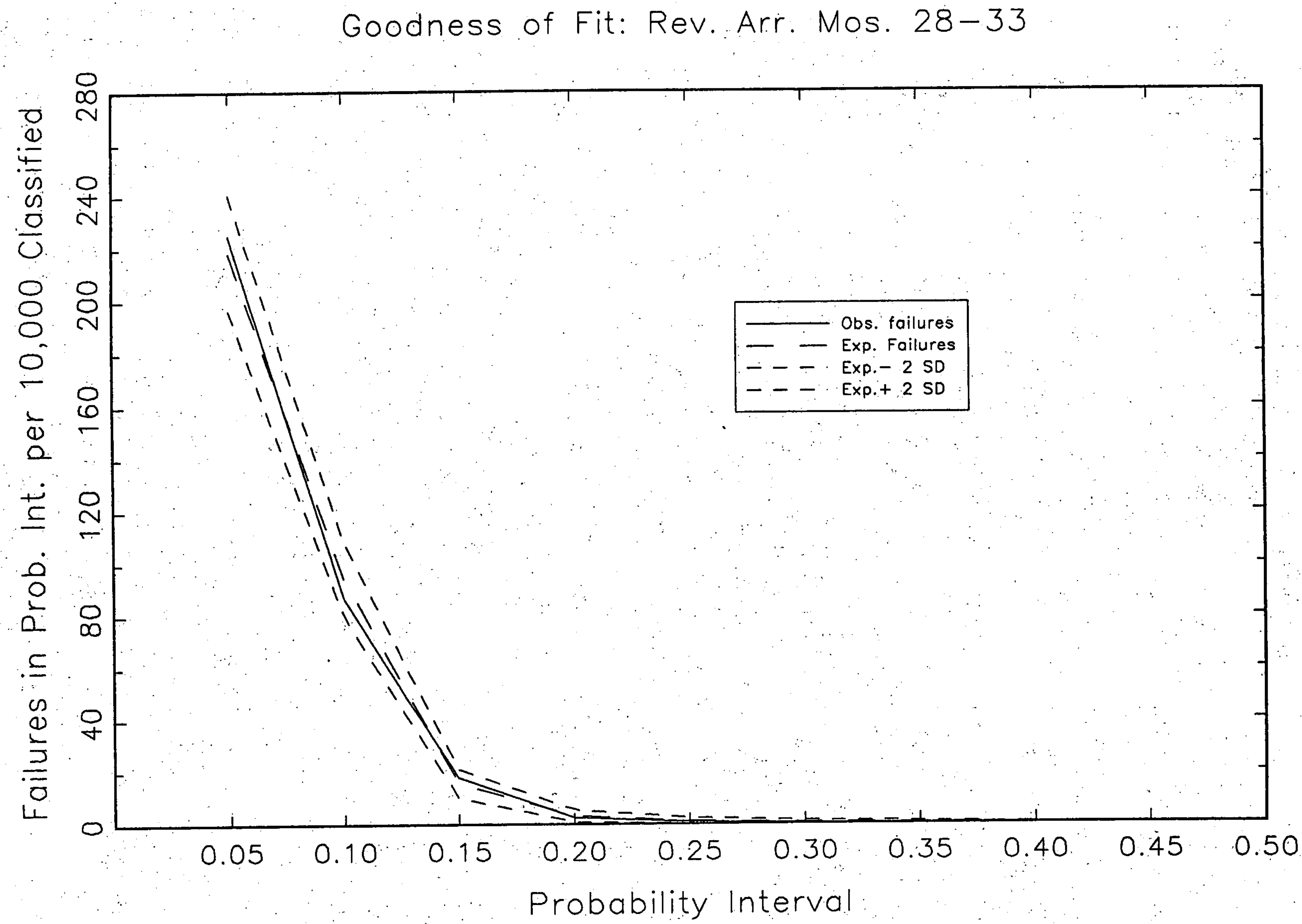
Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-2.27242	0.2343	-9.70	0.000
	2/6	-2.26091	0.2370	-9.54	0.000
	3/6	-2.42000	0.2317	-10.45	0.000
SEX	1/6	-0.48769	0.1280	-3.81	0.000
	2/6	-0.22145	0.1118	-1.98	0.048
	3/6	-0.24729	0.1054	-2.35	0.019
SPLIT	1/6	0.37391	0.1397	2.68	0.007
	2/6	-0.21462	0.1730	-1.24	0.215
	3/6	0.01011	0.1518	0.07	0.947
PRPRSN	1/6	0.29089	0.0633	4.59	0.000
	2/6	0.23155	0.0691	3.35	0.001
	3/6	0.10497	0.0755	1.39	0.165
ADMITS	1/6	0.13964	0.0340	4.11	0.000
	2/6	0.16823	0.0338	4.97	0.000
	3/6	0.11827	0.0337	3.51	0.000
VIOLENT	1/6	-0.05018	0.1201	-0.42	0.676
	2/6	0.05025	0.1200	0.42	0.675
	3/6	-0.44590	0.1194	-3.73	0.000
DRUG	1/6	0.11717	0.1071	1.09	0.274
	2/6	0.23016	0.1014	2.27	0.023
	3/6	-0.13724	0.0989	-1.39	0.165
OTHER	1/6	-0.00289	0.1585	-0.02	0.985
	2/6	-0.07988	0.1638	-0.49	0.626
	3/6	-0.29533	0.1486	-1.99	0.047
CIRCT2	1/6	-0.45493	0.3264	-1.39	0.163
	2/6	0.35762	0.3831	0.93	0.351
	3/6	0.25918	0.2542	1.02	0.308
CIRCT3	1/6	0.66100	0.3442	1.92	0.055
	2/6	0.35747	0.3070	1.16	0.244
	3/6	1.40492	0.3743	3.75	0.000
CIRCT5	1/6	-0.33124	0.2117	-1.56	0.118
	2/6	-0.65044	0.2099	-3.10	0.002
	3/6	-0.42197	0.1870	-2.26	0.024
CIRCT7	1/6	0.35342	0.3232	1.09	0.274
	2/6	-0.39554	0.3062	-1.29	0.196
	3/6	1.66992	0.3261	5.12	0.000
CIRCT8	1/6	0.69507	0.3565	1.95	0.051
	2/6	-0.02046	0.3407	-0.06	0.952
	3/6	0.62576	0.4459	1.40	0.161

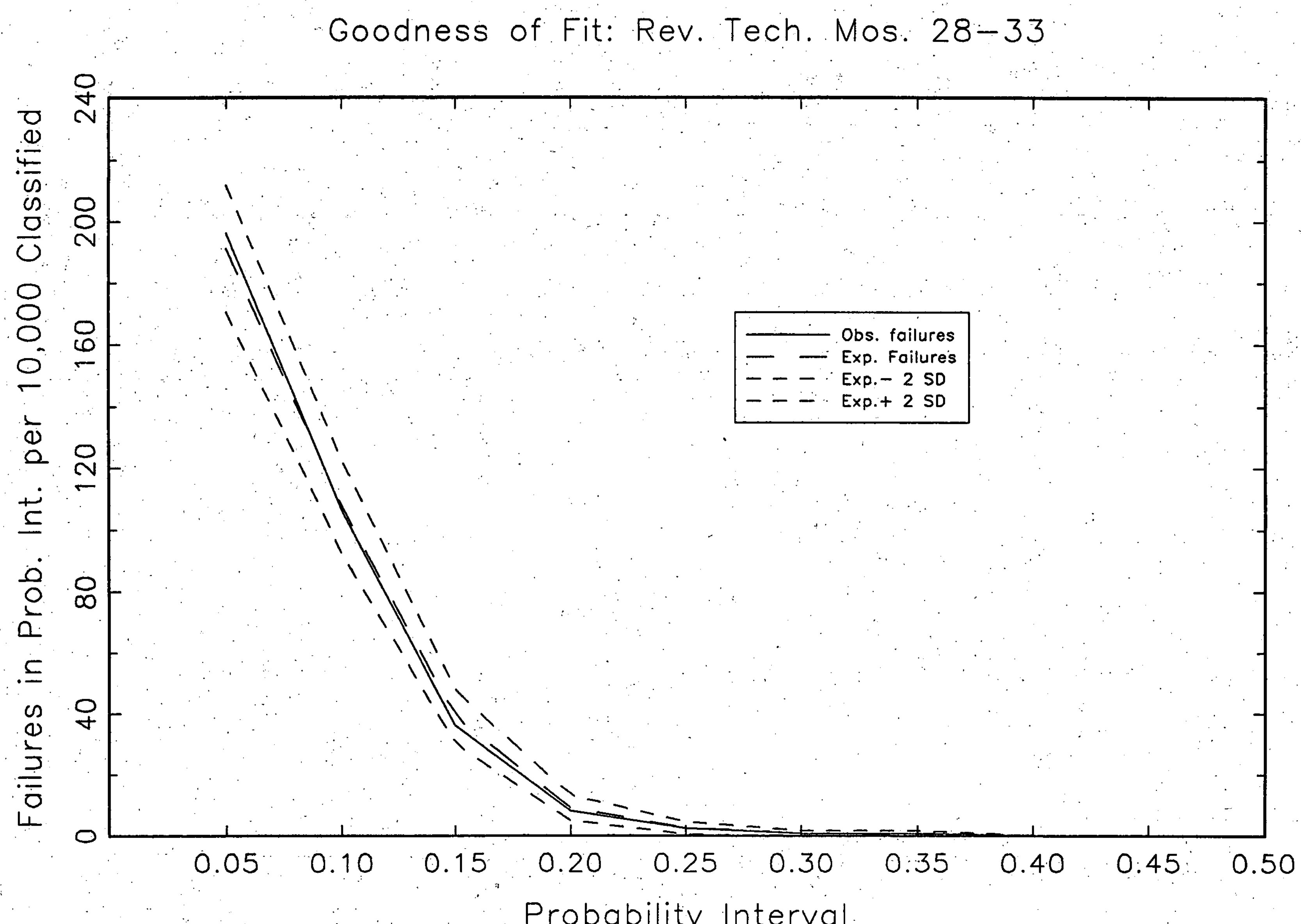
CIRCT10	1/6	-0.22084	0.2521	-0.88	0.381
	2/6	0.19461	0.2213	0.88	0.379
	3/6	0.03017	0.2402	0.13	0.900
CIRCT11	1/6	-0.12495	0.2145	-0.58	0.560
	2/6	-0.24215	0.2169	-1.12	0.264
	3/6	1.04975	0.2882	3.64	0.000
CIRCT12	1/6	0.29670	0.2721	1.09	0.276
	2/6	-0.53938	0.3406	-1.58	0.113
	3/6	-0.14683	0.3046	-0.48	0.630
CIRCT13	1/6	-0.30280	0.2213	-1.37	0.171
	2/6	-0.10700	0.2056	-0.52	0.603
	3/6	0.28367	0.1955	1.45	0.147
CIRCT14	1/6	0.00290	0.2949	0.01	0.992
	2/6	0.34752	0.4080	0.85	0.394
	3/6	0.42278	0.2562	1.65	0.099
CIRCT15	1/6	-0.21049	0.2746	-0.77	0.443
	2/6	0.08188	0.2504	0.33	0.744
	3/6	0.34430	0.4098	0.84	0.401
CIRCT16	1/6	0.36910	0.4425	0.83	0.404
	2/6	0.94349	0.3606	2.62	0.009
	3/6	2.04995	0.4215	4.86	0.000
CIRCT18	1/6	-0.50491	0.2271	-2.22	0.026
	2/6	-0.73237	0.2139	-3.42	0.001
	3/6	-0.46960	0.1889	-2.49	0.013
CIRCT19	1/6	0.29103	0.2494	1.17	0.243
	2/6	0.16429	0.2747	0.60	0.550
	3/6	1.68398	0.3122	5.39	0.000
CIRCT20	1/6	-0.43179	0.2713	-1.59	0.111
	2/6	-0.17490	0.2430	-0.72	0.472
	3/6	0.20119	0.2244	0.90	0.370
LAGEADM	1/6	-0.35842	0.0463	-7.75	0.000
	2/6	-0.26994	0.0464	-5.82	0.000
	3/6	-0.19224	0.0440	-4.37	0.000
REGION1	1/6	0.13705	0.2391	0.57	0.567
	2/6	-0.71584	0.3151	-2.27	0.023
	3/6	0.29385	0.2320	1.27	0.205
REGION2	1/6	-0.39065	0.2928	-1.33	0.182
	2/6	-0.05845	0.2327	-0.25	0.802
	3/6	-0.97215	0.3232	-3.01	0.003
REGION3	1/6	0.26328	0.1955	1.35	0.178
	2/6	0.47599	0.1800	2.64	0.008
	3/6	0.64287	0.1816	3.54	0.000
REGION4	1/6	-0.13054	0.1985	-0.66	0.511
	2/6	-0.33811	0.1915	-1.77	0.077
	3/6	-1.45178	0.2752	-5.27	0.000
LYRSUP2	1/6	-0.34948	0.1140	-3.07	0.002
	2/6	-0.53640	0.1164	-4.61	0.000
	3/6	-0.44053	0.1128	-3.91	0.000
LCOUNTS2	1/6	0.15696	0.0707	2.22	0.026
	2/6	0.04979	0.0776	0.64	0.521
	3/6	0.14819	0.0685	2.16	0.031
CATSUP	1/6	0.47970	0.1334	3.60	0.000
	2/6	0.86783	0.1260	6.89	0.000
	3/6	0.90084	0.1204	7.48	0.000

MEASURES OF FIT:

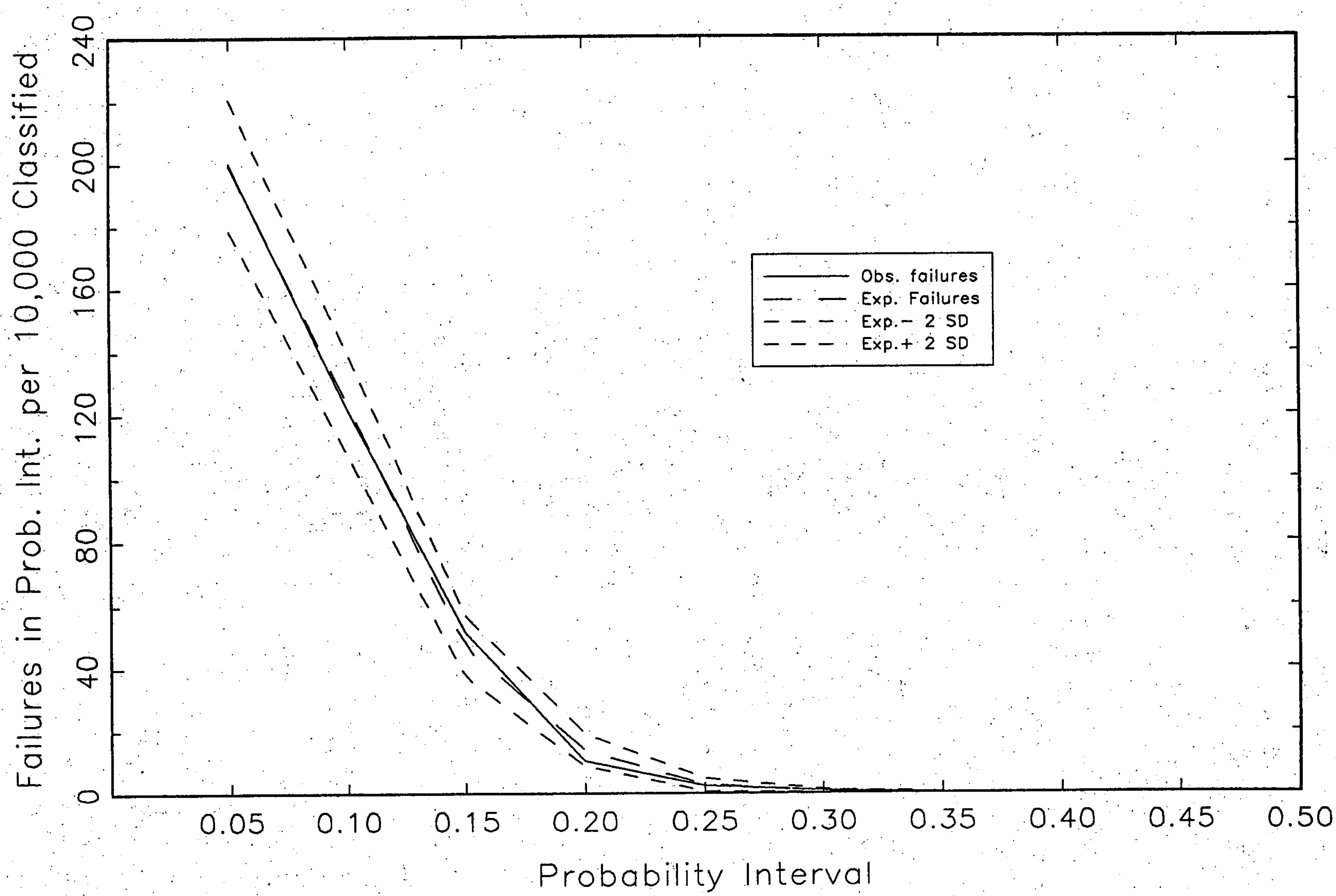
Test	LRX2	df	Prob
Overall	934.0767	90	0.000
CONSTANT	267.3632	3	0.000
SEX	22.4059	3	0.000
SPLIT	9.1356	3	0.028
PRPRSN	30.1882	3	0.000
ADMITS	45.2775	3	0.000
VIOLENT	14.3768	3	0.002
DRUG	8.5750	3	0.036
OTHER	4.1069	3	0.250
CIRCT2	4.0326	3	0.258
CIRCT3	17.8287	3	0.000
CIRCT5	15.5546	3	0.001
CIRCT7	29.4927	3	0.000
CIRCT8	5.5983	3	0.133
CIRCT10	1.6293	3	0.653
CIRCT11	15.3119	3	0.002
CIRCT12	4.1227	3	0.249
CIRCT13	4.4601	3	0.216
CIRCT14	3.3362	3	0.343
CIRCT15	1.4437	3	0.695
CIRCT16	28.5832	3	0.000
CIRCT18	20.5969	3	0.000
CIRCT19	30.0577	3	0.000
CIRCT20	3.9704	3	0.265
LAGEADM	102.7043	3	0.000
REGION1	7.3844	3	0.061
REGION2	10.5216	3	0.015
REGION3	19.4269	3	0.000
REGION4	30.3527	3	0.000
LYRSUP2	42.0501	3	0.000
LCOUNTS2	9.2832	3	0.026
CATSUP	104.9132	3	0.000

-2 Log Likelihood for full model: 15094.9406
-2 Log likelihood for restricted model: 16029.0173


Goodness of Fit: Rev. Arr. Mos. 28–33
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
280
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.– 2 SD
Exp.+ 2 SD


Goodness of Fit: Rev. Tech. Mos. 28–33
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.– 2 SD
Exp.+ 2 SD

Goodness of Fit: Absc. Mos. 28–33



Data Set: Months 34 through 39

N = 9852 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

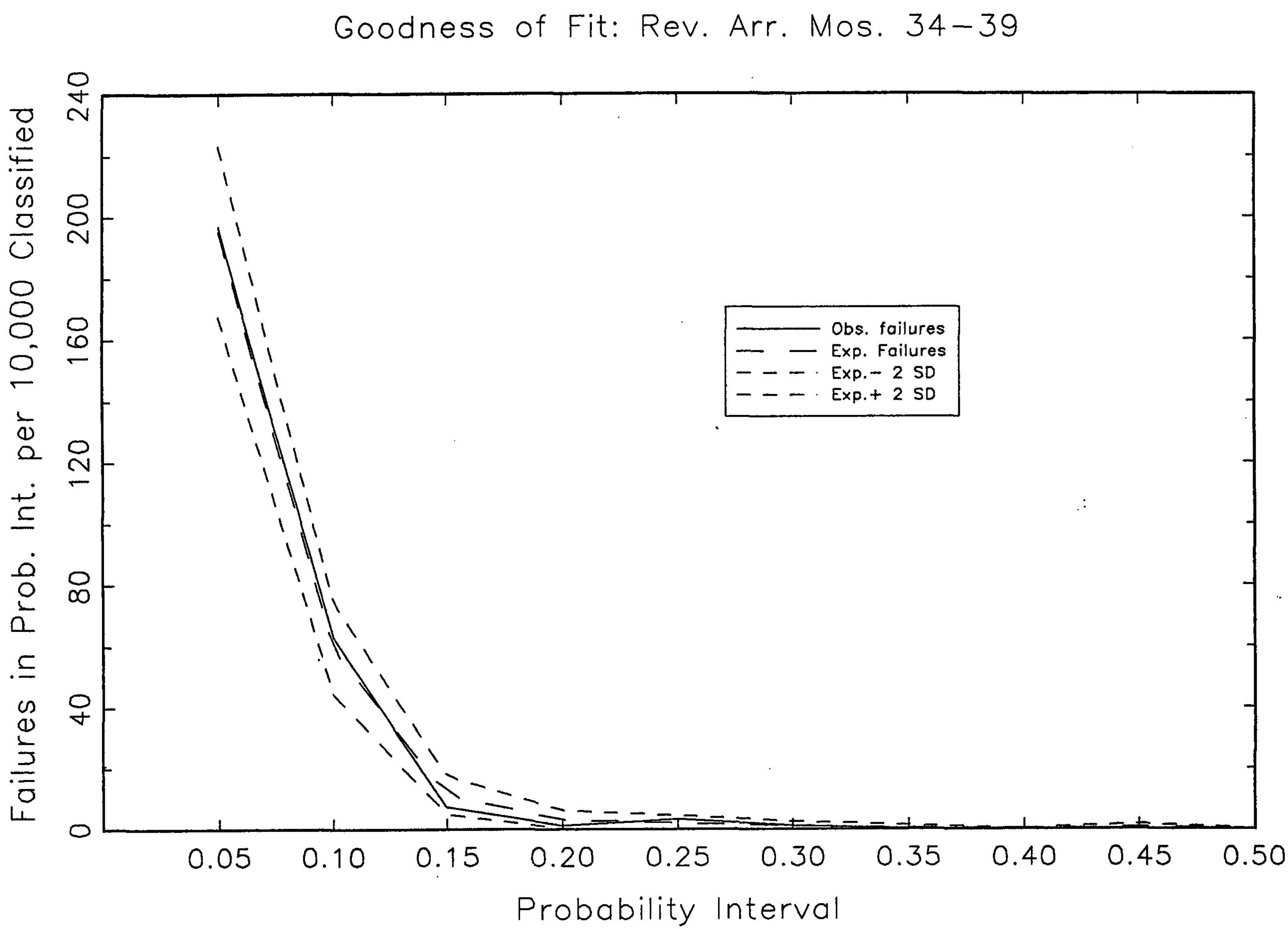
Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-1.32665	0.4138	-3.21	0.001
	2/6	-1.39657	0.3703	-3.77	0.000
	3/6	-1.05449	0.3740	-2.82	0.005
SEX	1/6	-0.64666	0.1969	-3.28	0.001
	2/6	-0.27506	0.1450	-1.90	0.058
	3/6	-0.28854	0.1312	-2.20	0.028
SPLIT	1/6	0.53593	0.1894	2.83	0.005
	2/6	0.28908	0.2011	1.44	0.151
	3/6	0.24040	0.1849	1.30	0.194
PRPRSN	1/6	0.31135	0.0911	3.42	0.001
	2/6	0.16515	0.1064	1.55	0.120
	3/6	0.31543	0.0849	3.71	0.000
ADMITS	1/6	0.20786	0.0458	4.53	0.000
	2/6	0.08374	0.0525	1.60	0.111
	3/6	0.04649	0.0474	0.98	0.326
VIOLENT	1/6	0.10248	0.1711	0.60	0.549
	2/6	-0.04763	0.1524	-0.31	0.755
	3/6	-0.56543	0.1513	-3.74	0.000
DRUG	1/6	0.11698	0.1620	0.72	0.470
	2/6	-0.01767	0.1371	-0.13	0.897
	3/6	-0.25189	0.1267	-1.99	0.047
OTHER	1/6	0.01204	0.2390	0.05	0.960
	2/6	-0.53743	0.2454	-2.19	0.029
	3/6	-0.24221	0.1806	-1.34	0.180
CIRCT2	1/6	0.34418	0.4085	0.84	0.399
	2/6	0.26477	0.4959	0.53	0.593
	3/6	0.28342	0.2841	1.00	0.318
CIRCT3	1/6	0.59025	0.6286	0.94	0.348
	2/6	0.15223	0.4633	0.33	0.742
	3/6	2.46039	0.8171	3.01	0.003
CIRCT5	1/6	-0.65312	0.3183	-2.05	0.040
	2/6	-0.21563	0.2576	-0.84	0.403
	3/6	-0.13360	0.2576	-0.52	0.604
CIRCT7	1/6	1.07848	0.5362	2.01	0.044
	2/6	-0.27016	0.4325	-0.62	0.532
	3/6	3.37980	0.7548	4.48	0.000
CIRCT8	1/6	-0.07829	0.8530	-0.09	0.927
	2/6	0.26228	0.4808	0.55	0.585
	3/6	3.05876	0.7942	3.85	0.000

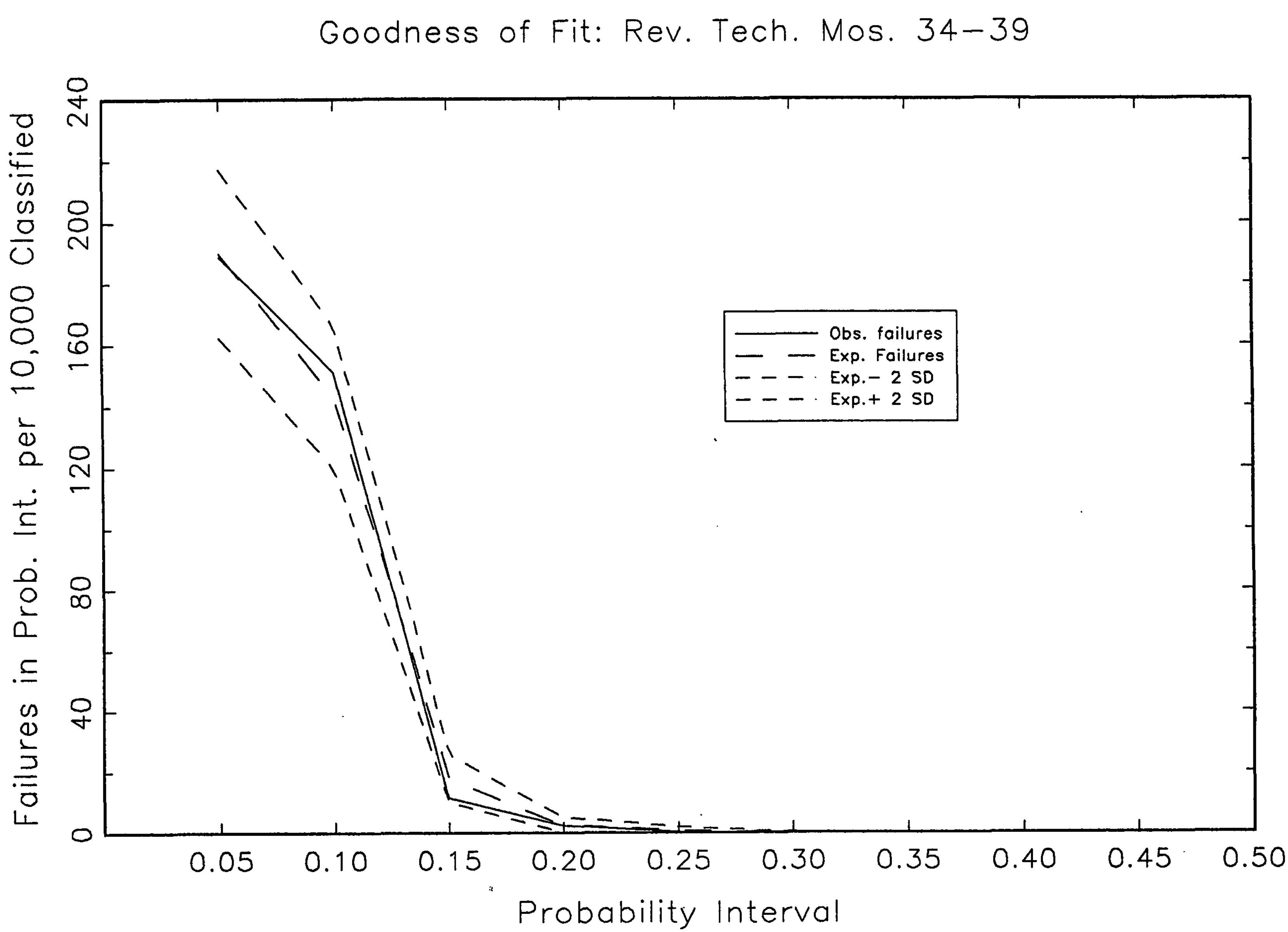
CIRCT10	1/6	-0.27976	0.3240	-0.86	0.388
	2/6	-0.24669	0.2643	-0.93	0.351
	3/6	0.11062	0.2854	0.39	0.698
CIRCT11	1/6	-0.35618	0.3785	-0.94	0.347
	2/6	-0.54754	0.3269	-1.67	0.094
	3/6	0.65406	0.3584	1.82	0.068
CIRCT12	1/6	-0.82788	0.5447	-1.52	0.129
	2/6	-0.26065	0.3612	-0.72	0.470
	3/6	0.00449	0.3900	0.01	0.991
CIRCT13	1/6	-1.08881	0.3340	-3.26	0.001
	2/6	-0.74197	0.2536	-2.93	0.003
	3/6	0.11829	0.2436	0.49	0.627
CIRCT14	1/6	-0.18154	0.4812	-0.38	0.706
	2/6	0.27455	0.4954	0.55	0.579
	3/6	0.17359	0.2942	0.59	0.555
CIRCT15	1/6	0.08616	0.3914	0.22	0.826
	2/6	-0.20706	0.3567	-0.58	0.562
	3/6	0.29797	0.4527	0.66	0.510
CIRCT16	1/6	-0.50203	1.0353	-0.48	0.628
	2/6	0.74433	0.5043	1.48	0.140
	3/6	1.35518	0.5922	2.29	0.022
CIRCT18	1/6	-0.13622	0.2924	-0.47	0.641
	2/6	-0.88525	0.3483	-2.54	0.011
	3/6	-0.03588	0.2695	-0.13	0.894
CIRCT19	1/6	0.60978	0.3548	1.72	0.086
	2/6	-0.10453	0.3698	-0.28	0.777
	3/6	1.99516	0.3335	5.98	0.000
CIRCT20	1/6	-0.17586	0.3172	-0.55	0.579
	2/6	-0.19704	0.2645	-0.75	0.456
	3/6	0.36961	0.2705	1.37	0.172
LAGEADM	1/6	-0.34070	0.0688	-4.95	0.000
	2/6	-0.16978	0.0619	-2.74	0.006
	3/6	-0.19614	0.0553	-3.55	0.000
REGION1	1/6	-0.34218	0.3393	-1.01	0.313
	2/6	-1.07842	0.3703	-2.91	0.004
	3/6	0.48962	0.2631	1.86	0.063
REGION2	1/6	-1.34967	0.5039	-2.68	0.007
	2/6	-0.83002	0.3250	-2.55	0.011
	3/6	-2.81381	0.7531	-3.74	0.000
REGION3	1/6	-0.01759	0.2639	-0.07	0.947
	2/6	-0.25925	0.2258	-1.15	0.251
	3/6	0.10554	0.2503	0.42	0.673
REGION4	1/6	-0.65090	0.2901	-2.24	0.025
	2/6	-0.68866	0.2360	-2.92	0.004
	3/6	-1.21182	0.3288	-3.69	0.000
LYRSUP2	1/6	-0.84117	0.2057	-4.09	0.000
	2/6	-0.81760	0.1786	-4.58	0.000
	3/6	-1.17748	0.1844	-6.39	0.000
LCOUNTS2	1/6	0.16816	0.1035	1.62	0.104
	2/6	-0.04623	0.1033	-0.45	0.655
	3/6	0.11302	0.0857	1.32	0.187
CATSUP	1/6	-0.22411	0.1875	-1.20	0.232
	2/6	0.40165	0.1702	2.36	0.018
	3/6	0.14437	0.1563	0.92	0.356

MEASURES OF FIT:

Test	LRX2	df	Prob
Overall	578.7555	90	0.000
CONSTANT	29.2002	3	0.000
SEX	18.0561	3	0.000
SPLIT	10.7112	3	0.013
PRPRSN	24.0217	3	0.000
ADMITS	22.2683	3	0.000
VIOLENT	14.5811	3	0.002
DRUG	4.6358	3	0.200
OTHER	6.3786	3	0.095
CIRCT2	1.7979	3	0.615
CIRCT3	9.8542	3	0.020
CIRCT5	4.9089	3	0.179
CIRCT7	24.3677	3	0.000
CIRCT8	15.0275	3	0.002
CIRCT10	1.7737	3	0.621
CIRCT11	7.2884	3	0.063
CIRCT12	2.7519	3	0.431
CIRCT13	19.0375	3	0.000
CIRCT14	0.8038	3	0.849
CIRCT15	0.8524	3	0.837
CIRCT16	7.2850	3	0.063
CIRCT18	6.5941	3	0.086
CIRCT19	38.1790	3	0.000
CIRCT20	2.9245	3	0.403
LAGEADM	40.8491	3	0.000
REGION1	13.5545	3	0.004
REGION2	26.1175	3	0.000
REGION3	1.5639	3	0.668
REGION4	24.9680	3	0.000
LYRSUP2	70.8617	3	0.000
LCOUNTS2	4.4877	3	0.213
CATSUP	8.0047	3	0.046

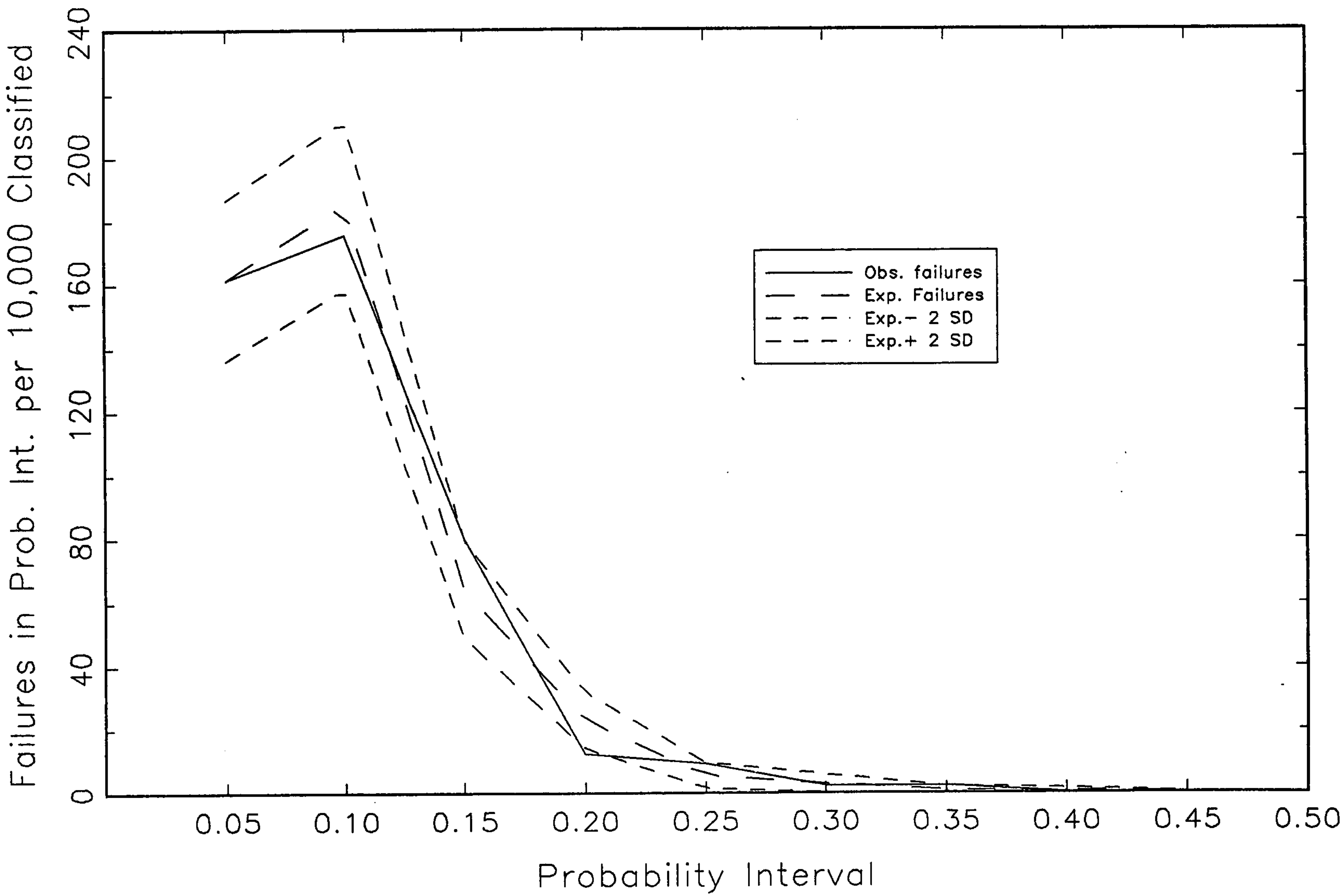
-2 Log Likelihood for full model: 8386.2633
-2 Log likelihood for restricted model: 8965.0188


Goodness of Fit: Rev. Arr. Mos. 34-39
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD



Goodness of Fit: Rev. Tech. Mos. 34-39
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD

Goodness of Fit: Absc. Mos. 34–39



Data Set: Months 40 through 48

N = 2419 cases

DEPENDENT CATEORIES ARE DESIGNATED AS:

1 - Revoke arrest
2 - Revoke Tech.
3 - Abscond
6 - Release or still active

Tolerance of 0.0000

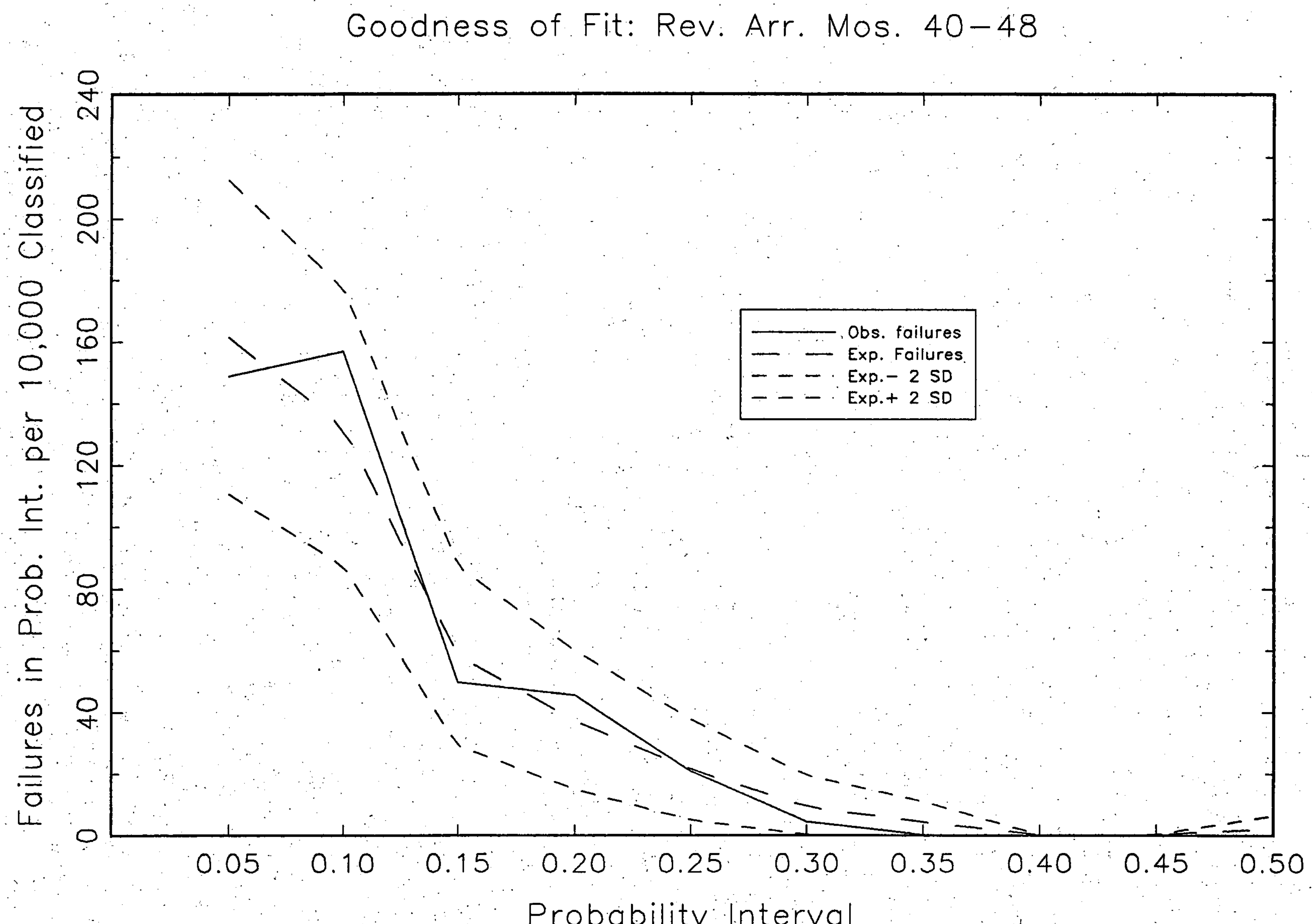
Variable	Comparison	Logit Estimate	Std. Error	t-value	p>|t|
CONSTANT	1/6	-0.53268	0.6724	-0.79	0.428
	2/6	-0.77227	0.6454	-1.20	0.231
	3/6	-1.32411	0.6881	-1.92	0.054
SEX	1/6	-0.27950	0.2978	-0.94	0.348
	2/6	-0.27776	0.2507	-1.11	0.268
	3/6	-0.53418	0.2413	-2.21	0.027
SPLIT	1/6	1.18796	0.2913	4.08	0.000
	2/6	0.47125	0.3240	1.45	0.146
	3/6	0.46899	0.2783	1.69	0.092
PRPRSN	1/6	0.21690	0.1860	1.17	0.243
	2/6	0.09464	0.1664	0.57	0.569
	3/6	0.21409	0.1507	1.42	0.156
ADMITS	1/6	0.12090	0.0902	1.34	0.180
	2/6	0.13199	0.0749	1.76	0.078
	3/6	0.11482	0.0671	1.71	0.087
VIOLENT	1/6	0.13208	0.2760	0.48	0.632
	2/6	-0.14832	0.2671	-0.56	0.579
	3/6	-0.53110	0.2347	-2.26	0.024
DRUG	1/6	0.20358	0.2815	0.72	0.470
	2/6	0.29532	0.2311	1.28	0.201
	3/6	-0.13087	0.2209	-0.59	0.554
OTHER	1/6	0.29098	0.3827	0.76	0.447
	2/6	0.15602	0.3460	0.45	0.652
	3/6	0.10454	0.2954	0.35	0.723
CIRCT2	1/6	-1.12929	0.8509	-1.33	0.184
	2/6	-0.70337	1.2451	-0.56	0.572
	3/6	0.52415	0.6258	0.84	0.402
CIRCT3	1/6	0.28700	0.8063	0.36	0.722
	2/6	1.17780	0.7026	1.68	0.094
	3/6	0.80596	1.4427	0.56	0.576
CIRCT5	1/6	-0.95463	0.6050	-1.58	0.115
	2/6	-0.29099	0.4790	-0.61	0.543
	3/6	-0.77660	0.3574	-2.17	0.030
CIRCT7	1/6	0.70884	0.6693	1.06	0.290
	2/6	0.68654	0.6945	0.99	0.323
	3/6	3.39708	1.0575	3.21	0.001
CIRCT8	1/6	-0.45035	1.1453	-0.39	0.694
	2/6	-0.33676	1.1333	-0.30	0.766
	3/6	2.92704	1.1025	2.65	0.008

CIRCT10	1/6	-1.23955	0.5435	-2.28	0.023
	2/6	-0.02485	0.4613	-0.05	0.957
	3/6	0.21567	0.6281	0.34	0.731
CIRCT11	1/6	-0.82919	0.5414	-1.53	0.126
	2/6	-0.18000	0.4352	-0.41	0.679
	3/6	0.70584	0.5901	1.20	0.232
CIRCT12	1/6	-1.34494	0.7776	-1.73	0.084
	2/6	-0.67060	0.6766	-0.99	0.322
	3/6	0.43394	0.6980	0.62	0.534
CIRCT13	1/6	-2.49799	0.7696	-3.25	0.001
	2/6	-0.27742	0.4311	-0.64	0.520
	3/6	0.72211	0.5383	1.34	0.180
CIRCT14	1/6	0.24414	0.6256	0.39	0.696
	2/6	1.26353	0.8683	1.46	0.146
	3/6	0.80353	0.6276	1.28	0.200
CIRCT15	1/6	-0.95378	0.7764	-1.23	0.219
	2/6	-0.62249	0.6590	-0.94	0.345
	3/6	-0.07176	0.8553	-0.08	0.933
CIRCT18	1/6	-0.74595	0.5583	-1.34	0.182
	2/6	-0.34586	0.4946	-0.70	0.484
	3/6	-0.37503	0.3219	-1.17	0.244
CTS16_19	1/6	-1.30326	0.7768	-1.68	0.093
	2/6	0.04294	0.5129	0.08	0.933
	3/6	0.76627	0.6571	1.17	0.244
CIRCT20	1/6	-1.29702	0.5874	-2.21	0.027
	2/6	-0.47954	0.5295	-0.91	0.365
	3/6	0.76610	0.5816	1.32	0.188
LAGEADM	1/6	-0.59094	0.1139	-5.19	0.000
	2/6	-0.29044	0.1054	-2.76	0.006
	3/6	-0.32892	0.0921	-3.57	0.000
REGION1	1/6	-0.51507	0.5221	-0.99	0.324
	2/6	-1.24157	0.7862	-1.58	0.114
	3/6	0.37394	0.6583	0.57	0.570
REGION2	1/6	-1.52972	0.6004	-2.55	0.011
	2/6	-1.39933	0.6008	-2.33	0.020
	3/6	-1.91526	1.1226	-1.71	0.088
REGION3	1/6	-0.59597	0.4150	-1.44	0.151
	2/6	-0.20649	0.4289	-0.48	0.630
	3/6	1.61936	0.5034	3.22	0.001
REGION4	1/6	-0.56820	0.3990	-1.42	0.154
	2/6	-0.43404	0.4192	-1.04	0.301
	3/6	-0.56958	0.6522	-0.87	0.383
LYRSUP2	1/6	-0.34722	0.3026	-1.15	0.251
	2/6	-0.85276	0.2731	-3.12	0.002
	3/6	-0.73629	0.2614	-2.82	0.005
LCOUNTS2	1/6	-0.20163	0.1865	-1.08	0.280
	2/6	0.04012	0.1594	0.25	0.801
	3/6	0.21247	0.1307	1.63	0.104
CATSUP	1/6	0.45313	0.3049	1.49	0.137
	2/6	0.60651	0.2708	2.24	0.025
	3/6	0.13924	0.2398	0.58	0.561

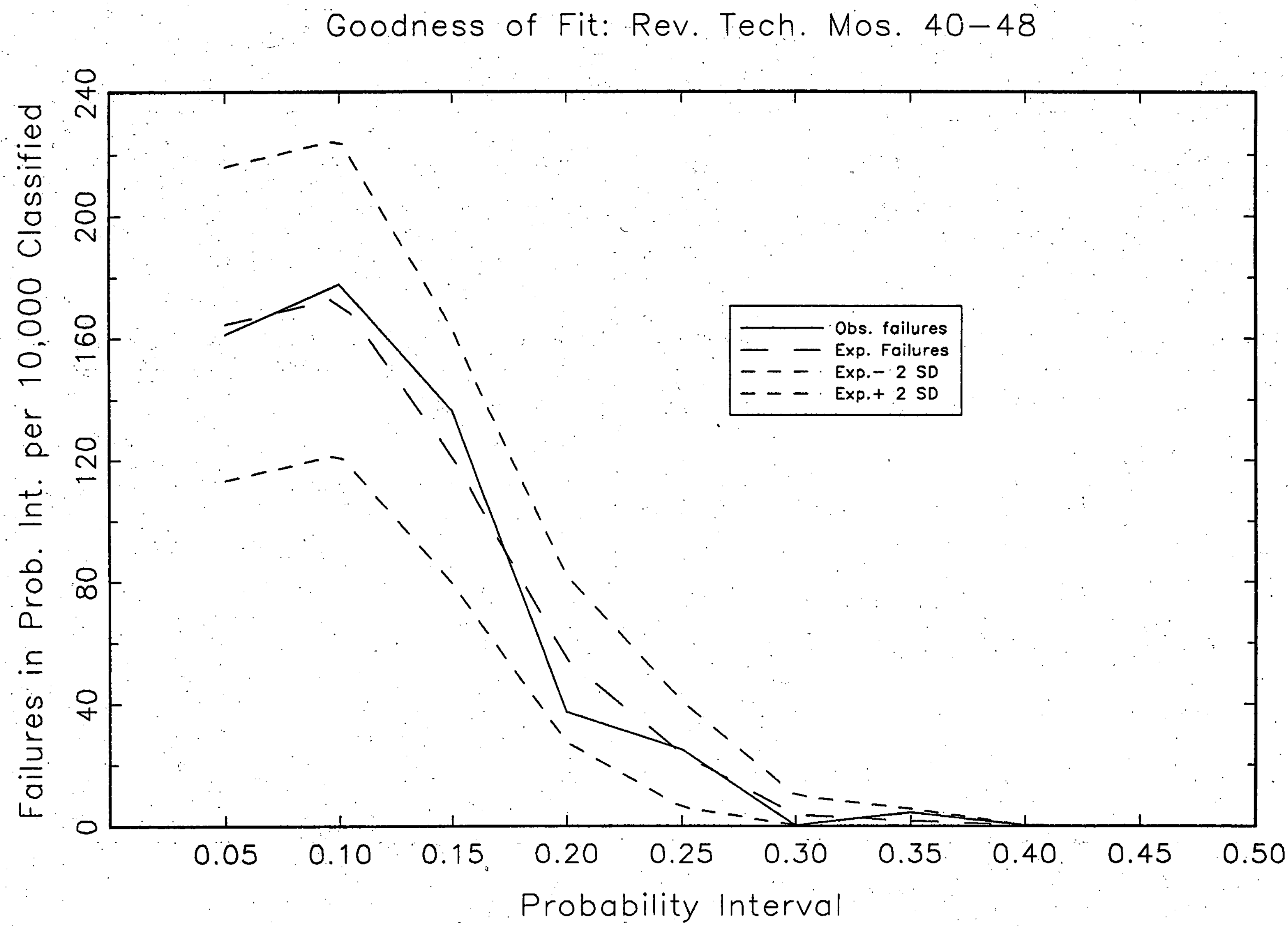
MEASURES OF FIT:

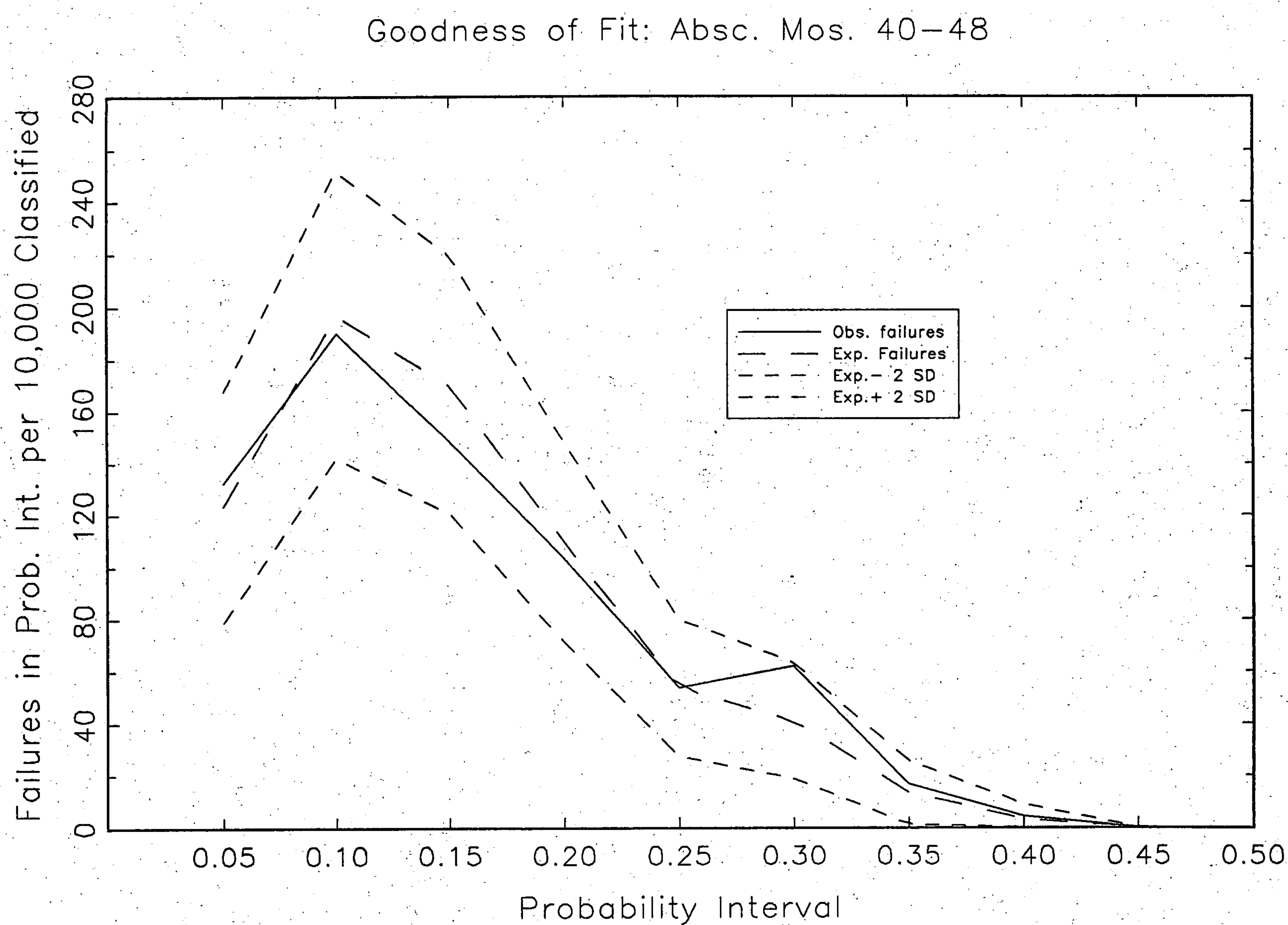
Test	LRX2	df	Prob
Overall	305.7487	87	0.000
CONSTANT	5.1137	3	0.164
SEX	6.4215	3	0.093
SPLIT	18.9818	3	0.000
PRPRSN	3.1734	3	0.366
ADMITS	6.1374	3	0.105
VIOLENT	5.7332	3	0.125
DRUG	2.6000	3	0.457
OTHER	0.8009	3	0.849
CIRCT2	2.9490	3	0.400
CIRCT3	3.1009	3	0.376
CIRCT5	6.6762	3	0.083
CIRCT7	11.8651	3	0.008
CIRCT8	7.4150	3	0.060
CIRCT10	5.4767	3	0.140
CIRCT11	4.1471	3	0.246
CIRCT12	4.3920	3	0.222
CIRCT13	13.0895	3	0.004
CIRCT14	3.4940	3	0.322
CIRCT15	2.2921	3	0.514
CIRCT18	3.0858	3	0.379
CTS16_19	4.3734	3	0.224
CIRCT20	7.7595	3	0.051
LAGEADM	40.7249	3	0.000
REGION1	3.7926	3	0.285
REGION2	13.3903	3	0.004
REGION3	13.7619	3	0.003
REGION4	3.2693	3	0.352
LYRSUP2	17.0583	3	0.001
LCOUNTS2	4.1373	3	0.247
CATSUP	6.8783	3	0.076

-2 Log Likelihood for full model: 2757.5275
-2 Log likelihood for restricted model: 3063.2762



Goodness of Fit: Rev. Arr. Mos. 40–48
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.– 2 SD
Exp.+ 2 SD


Goodness of Fit: Rev. Tech. Mos. 40-48
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD



Goodness of Fit: Absc. Mos. 40-48
Failures in Prob. Int. per 10,000 Classified
0
40
80
120
160
200
240
280
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Probability Interval
Obs. failures
Exp. Failures
Exp.- 2 SD
Exp.+ 2 SD