AD-A062 728 DESMATICS INC STATE COLLEGE PA F/G 12/1

STATISTICAL INFERENCE PROCEDURES FOR A LOGISTIC IMPACT ACCELERA--ETC(U)

DEC 78 J J PETERSON, D E SMITH N00014-74-C-0154

UNCLASSIFIED TR-102-7 NL

1 OF 1

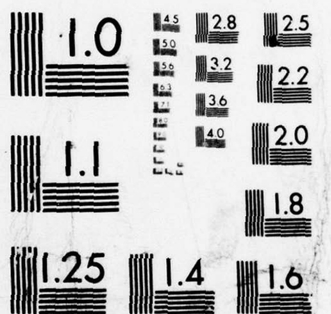
AD A062728

END

DATE FILMED

3 -- 79

DDC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

LEVEL II

13

— STATISTICS —

— OPERATIONS RESEARCH —

— MATHEMATICS —

391156

DESMATICS, INC.

P.O. Box 618

LEVEL II

DESMATICS, INC.

P. O. Box 618
State College, Pa. 16801
Phone: (814) 238-9621

Applied Research in Statistics - Mathematics - Operations Research

(13)

(6)

STATISTICAL INFERENCE PROCEDURES FOR A LOGISTIC IMPACT ACCELERATION INJURY PREDICTION MODEL

BY

(10) John J. Peterson

[redacted]

Dennis E. Smith

(12) 35p.

(14) TR-1Ø2-7

(9) TECHNICAL REPORT NO. 102-7

(11) Decem[redacted]978

This study was supported by the Office of Naval Research under Contract No. (15) NØØ014-74-C-Ø154, Task No. NR 207-037

391 156

78 12 13 017

TABLE OF CONTENTS

I. INTRODUCTION

The forces occurring in a naval aircraft ditching accident often cause injuries or fatalities to the aircraft occupants. Impact of the aircraft on the water can produce serious injury, especially to the neck and head, causing the pilot to lose consciousness. Unfortunately, fatal consequences may result from such an occurrence. Thus, it is important to be aware of how the dynamic and physical variables involved in a ditching accident relate to the likelihood of pilot injury.

The U.S. Navy's impact acceleration research program being conducted by the Naval Aerospace Medical Research Laboratory (NAMRL) Detachment is engaged in experimentation concerning dynamic response of the human and simian head/neck system as a function of motion and anthropometric parameters. The development of a model based primarily on the information from the experimental data allows important statistical inferences to be made concerning injury probability as a function of dynamic and physical variables. Two previous technical reports [2, 4] have discussed model formulation and accuracy, respectively.

This report addresses itself to the extraction of important inferences using a logistic function to model injury likelihood. As an example, inferences are made for a logistic model constructed from the data obtained in the Monte Carlo simulation study described in a recent technical report [4]. Specific information on how the data was simulated can be obtained in that report. Since the data was simulated, the true underlying model parameters are known. Therefore, this data base allows an assessment of the accuracy of the inference procedures derived. The inference of foremost importance in this study is the estimation of threshold levels which allow for only a small prespecified chance of head/neck injury.

II. BACKGROUND

The statistical impact acceleration injury models, described in the previous technical reports [2, 3, 4], may be used to predict the probability of head/neck injury for a given set of anthropometric and dynamic response variables. Due to sampling variability in model parameter estimates, the prediction of the probability of an injury for a given set of conditions has certain degree of error variability associated with it. Inference procedures that account for this variability are used to estimate "safe" threshold levels, i.e., threshold levels that are reliable in keeping injury probability to an acceptable minimum.

The logistic injury prediction model is of the form

$$P(\underline{x}) = \{1 + \exp[-(\underline{x}'\underline{\beta})]\}^{-1}$$

where $\underline{x} = (1, x_1, x_2, \ldots, x_k)$ denotes the set of independent variables considered

and $\underline{\beta} = (\beta_0, \beta_1, \beta_2, \ldots, \beta_k)$ denotes the set of parameter values.

If $\hat{\underline{\beta}}$ denotes the maximum likelihood estimate of $\underline{\beta}$, then $\hat{P}(\underline{x}) = \{1 + \exp[-(\underline{x}'\hat{\underline{\beta}})]\}^{-1}$ is an estimate of $P(\underline{x})$. To derive the inference procedures concerning $\underline{\beta}$ and $P(\underline{x})$ that give consideration to the variability of $\hat{\underline{\beta}}$, the sampling distribution of $\hat{\underline{\beta}}$ must be known. Because $\hat{\underline{\beta}}$ is a function of the data, it is the outcome of a random vector.

Choosing $\hat{\underline{\beta}}$ to be the maximum likelihood estimate of $\underline{\beta}$ allows exploitation of the rich theory of maximum likelihood estimation to obtain an approximation to the distribution of $\hat{\underline{\beta}}$. It can be shown, using this theory, that $\hat{\underline{\beta}}$ is asymptotically distributed as multivariate normal with mean vector $\underline{\beta}$ and a certain variance-covariance matrix, $\underline{\Sigma}$. The inherent variability associated with inferences concerning $\underline{\beta}$ and $P(\underline{x})$ are directly related to the variance-covariance structure found in the $\underline{\Sigma}$ matrix. This matrix, unfortunately, is an unknown parameter of the distribution of $\hat{\underline{\beta}}$. However, by applying maximum likelihood theory to the logistic model, it is possible to show that $\underline{\Sigma}$ is a function of $\underline{\beta}$. Thus, $\underline{\Sigma}$ can be estimated by the substitution of $\hat{\underline{\beta}}$ for $\underline{\beta}$.

The FORTRAN subroutine of Jones [1] was used to compute $\hat{\underline{\beta}}$ and $\hat{\underline{\Sigma}}$ (the estimate of $\underline{\Sigma}$) from the simulated data. The two statistics $\hat{\underline{\beta}}$ and $\hat{\underline{\Sigma}}$ are sufficient to form any (approximate) inferences about $\underline{\beta}$ and $P(\underline{x})$ that may be desired. It should be noted that as the amount of data increases, the estimates, $\hat{\underline{\beta}}$ and $\hat{\underline{\Sigma}}$, get closer to the true values of the parameters $\underline{\beta}$ and $\underline{\Sigma}$, and the sampling distribution of $\hat{\underline{\beta}}$ becomes closer to multivariate normal.

Estimates of $\underline{\beta}$ and $\underline{\Sigma}$ for models A and B (defined in a previous report [4]) are given in Figures 1 through 4. Sample sizes of $N = 100$ and $N = 1000$ were used for both models.

$\underline{\beta}$ = (0.00, -0.25, 0.50, -0.75, 1.00, -1.25, 1.50)

$\hat{\underline{\beta}}$ = (-0.43, -0.43, 0.92, -0.78, 1.53, -1.03, 0.89)

$\hat{\underline{\Sigma}}$ =

0.64E-01	0.17E-01	-0.38E-02	-0.81E-02	-0.23E-01	0.16E-01	0.48E-02
0.17E-01	0.20E 00	-0.15E-01	0.20E-01	-0.46E-01	-0.10E-01	0.74E-04
-0.38E-02	-0.15E 00	0.18E 00	-0.16E-01	0.46E-01	-0.18E-01	0.15E-01
-0.81E-02	0.20E-01	-0.16E-01	0.20E 00	-0.51E-01	0.10E-01	0.25E-01
-0.23E-01	-0.46E-01	0.46E-01	-0.51E-01	0.22E 00	-0.16E-01	0.44E-01
0.16E-01	-0.10E-01	-0.18E-01	0.10E-01	-0.16E-01	0.17E 00	0.13E-04
0.48E-02	0.74E-04	0.15E-01	0.25E-01	0.44E-01	0.13E-04	0.22E-00

Figure 1: Estimates of $\underline{\beta}$ and $\underline{\Sigma}$ for Model A Based on a Sample Size of N = 100

$\underline{\beta}$ = (0.00, -0.25, 0.50, -0.75, 1.00, -1.25, 1.50)

$\hat{\underline{\beta}}$ = (-0.07, -0.19, 0.56, -0.80, 1.02, -1.27, 1.49)

$$\hat{\underline{\Sigma}} = \begin{bmatrix} 0.56E-02 & -0.82E-04 & -0.21E-03 & -0.22E-03 & -0.77E-04 & 0.27E-03 & -0.34E-03 \\ -0.82E-04 & 0.17E-01 & -0.18E-02 & 0.73E-03 & -0.64E-03 & 0.58E-03 & -0.22E-03 \\ -0.21E-03 & -0.18E-02 & 0.17E-01 & -0.73E-03 & 0.25E-02 & -0.16E-02 & 0.27E-02 \\ -0.22E-03 & 0.73E-03 & -0.73E-03 & 0.18E-01 & -0.14E-02 & 0.20E-02 & -0.15E-02 \\ -0.77E-04 & -0.64E-03 & 0.25E-02 & -0.14E-02 & 0.18E-01 & -0.27E-02 & 0.36E-02 \\ 0.27E-03 & 0.58E-03 & -0.16E-02 & 0.20E-02 & -0.27E-02 & 0.19E-01 & -0.39E-02 \\ -0.34E-03 & -0.22E-03 & 0.27E-02 & -0.15E-02 & 0.36E-02 & -0.39E-02 & 0.20E-01 \end{bmatrix}$$

Figure 2: Estimates of $\underline{\beta}$ and $\underline{\Sigma}$ for Model A Based on a Sample Size of N = 1000

$\underline{\beta}$ = (-2.00, -0.25, 0.50, -0.75, 1.00, -1.25, 1.50)

$\hat{\underline{\beta}}$ = (-2.27, -0.06, 0.64 -0.00, 0.14, -1.85, 2.31)

$$\hat{\underline{\Sigma}} = \begin{bmatrix} 0.21\text{E } 00 & 0.22\text{E}{-01} & -0.56\text{E}{-02} & -0.22\text{E}{-01} & -0.13\text{E } 00 & 0.15\text{E } 00 & -0.18\text{E } 00 \\ 0.22\text{E}{-01} & 0.33\text{E } 00 & -0.33\text{E}{-01} & -0.36\text{E}{-02} & -0.51\text{E}{-01} & -0.34\text{E}{-01} & 0.30\text{E}{-01} \\ -0.56\text{E}{-01} & -0.33\text{E}{-01} & 0.33\text{E } 00 & 0.12\text{E}{-01} & 0.53\text{E}{-01} & -0.54\text{E}{-02} & -0.67\text{E}{-01} \\ -0.22\text{E}{-01} & -0.36\text{E}{-02} & 0.12\text{E}{-01} & 0.31\text{E } 00 & -0.42\text{E}{-01} & 0.19\text{E}{-01} & 0.86\text{E}{-01} \\ -0.13\text{E } 00 & -0.51\text{E}{-01} & 0.53\text{E}{-01} & -0.42\text{E}{-01} & 0.37\text{E } 00 & -0.65\text{E } 00 & 0.13\text{E } 00 \\ 0.15\text{E } 00 & -0.34\text{E}{-01} & -0.54\text{E}{-02} & 0.19\text{E}{-01} & -0.65\text{E}{-01} & 0.39\text{E } 00 & -0.10\text{E } 00 \\ -0.18\text{E } 00 & 0.30\text{E}{-01} & -0.67\text{E}{-01} & 0.86\text{E}{-01} & 0.13\text{E } 00 & -0.10\text{E } 00 & 0.56\text{E } 00 \end{bmatrix}$$

Figure 3: Estimates of $\underline{\beta}$ and $\underline{\Sigma}$ for Model B Based on a Sample Size of N = 100

$$\underline{\beta} = (-2.00,\ -0.25,\ 0.50,\ -0.75,\ 1.00,\ -1.25,\ 1.50)$$

$$\hat{\underline{\beta}} = (-2.15,\ -0.26,\ 0.54,\ -0.85,\ 0.91,\ -1.45,\ 1.81)$$

$$\hat{\underline{\Sigma}} = \begin{bmatrix} 0.17E-01 & 0.18E-02 & -0.37E-02 & 0.50E-02 & -0.58E-02 & 0.10E-01 & -0.13E-01 \\ 0.18E-02 & 0.28E-01 & -0.16E-02 & 0.21E-02 & -0.82E-03 & 0.22E-02 & -0.16E-02 \\ -0.37E-02 & -0.16E-02 & 0.27E-01 & -0.67E-03 & 0.28E-02 & -0.30E-02 & 0.29E-02 \\ 0.50E-02 & 0.21E-02 & -0.67E-03 & 0.30E-01 & -0.54E-03 & 0.29E-02 & -0.13E-02 \\ -0.58E-02 & -0.83E-03 & 0.28E-02 & -0.54E-03 & 0.29E-01 & -0.38E-02 & 0.39E-02 \\ 0.10E-01 & 0.22E-02 & -0.30E-02 & 0.29E-02 & -0.38E-02 & 0.33E-01 & -0.74E-02 \\ -0.13E-01 & -0.16E-02 & 0.29E-02 & -0.13E-02 & 0.39E-02 & -0.74E-02 & 0.37E-01 \end{bmatrix}$$

Figure 4: Estimates of $\underline{\beta}$ and $\underline{\Sigma}$ for Model B Based on a Sample Size of N = 1000

III. INTERVAL ESTIMATION AND HYPOTHESIS TESTING

Point estimation of $\underline{\beta}$ and $P(\underline{x})$ was discussed in previous technical reports [2, 4]. This report is concerned with inferences about $\underline{\beta}$ and $P(\underline{x})$ that account for sampling variability in $\hat{\underline{\beta}}$. Point estimates do not do this. Confidence interval estimates and tests of hypotheses concerning $\underline{\beta}$ and $P(\underline{x})$ are sensitive to the variability of $\hat{\underline{\beta}}$, so attention is directed to them. Such statistical structures are important in the development of safe threshold levels.

A. INTERVAL ESTIMATION

In this section confidence intervals are derived for β_i and for $P(\underline{x})$, for a given set of independent variables $\underline{x}$. The fact that $\hat{\underline{\beta}}$ is asymptotically multivariate normal with mean vector $\underline{\beta}$ and covariance matrix $\underline{\Sigma}$ enables the derivation of an approximate $1 - \alpha$ confidence interval for $\underline{x}'\underline{\beta}$ and for $P(\underline{x})$ corresponding to a given $\underline{x}$ vector. Since, asymptotically,

$$\hat{\beta} \sim N(\underline{\beta}, \hat{\underline{\Sigma}})$$

it follows that, asymptotically,

$$\underline{x}'\hat{\underline{\beta}} \sim N(\underline{x}'\underline{\beta}, q^2(\underline{x}))$$

where $q(\underline{x}) = (\underline{x}'\hat{\underline{\Sigma}}\underline{x})^{\frac{1}{2}}$. This in turn implies that, approximately,

$$\text{Prob}(\underline{x}'\hat{\underline{\beta}} - z_{\alpha/2}q(\underline{x}) \leq \underline{x}'\underline{\beta} \leq \underline{x}'\hat{\underline{\beta}} + z_{\alpha/2}q(\underline{x})) = 1 - \alpha,$$

where z_α is the upper α point of a standard normal distribution. Thus, the outcome of this random interval is an approximate $1 - \alpha$ confidence interval for $\underline{x}'\underline{\beta}$.

The event

$$\{\underline{x}'\hat{\underline{\beta}} - z_{\alpha/2}q(\underline{x}) \leq \underline{x}'\underline{\beta} \leq \underline{x}'\hat{\underline{\beta}} + z_{\alpha/2}q(\underline{x})\}$$

is equivalent to the event

$$E = \{(1 + \exp[-(\underline{x}'\hat{\underline{\beta}} - z_{\alpha/2}q(\underline{x}))])^{-1} \leq P(\underline{x}) \leq (1 + \exp[-(\underline{x}'\hat{\underline{\beta}} + z_{\alpha/2}q(\underline{x}))])^{-1}\}.$$

Therefore, the probability of E is also $1 - \alpha$, approximately. Thus, the random interval

$$[(1 + \exp[-(\underline{x}'\hat{\underline{\beta}} - z_{\alpha/2}q(\underline{x}))])^{-1},\ (1 + \exp[-(\underline{x}'\hat{\underline{\beta}} + z_{\alpha/2}q(\underline{x}))])^{-1}] \quad (1)$$

contains $P(\underline{x})$ with probability approximately $1 - \alpha$. The outcome of this random interval is then an approximate $1 - \alpha$ confidence interval for $P(\underline{x})$.

It is also possible to construct confidence intervals for a given parameter β_i by using $\hat{\underline{\beta}}$ and $\hat{\underline{\Sigma}}$. Since $\hat{\beta}_i$ is a maximum likelihood estimate for β_i, $\hat{\beta}_i$ is asymptotically normal with mean β_i and variance v_i^2, where v_i^2 is the $(i, i)^{th}$ element of $\hat{\underline{\Sigma}}$. This fact implies that

$$\text{Prob}(\hat{\beta}_i - z_{\alpha/2}v_i \leq \beta_i \leq \hat{\beta}_i + z_{\alpha/2}v_i) = 1 - \alpha, \text{ approximately.}$$

Thus, the random interval

$$(\hat{\beta}_i - z_{\alpha/2}v_i,\ \hat{\beta}_i + z_{\alpha/2}v_i) \quad (2)$$

contains β_i with probability approximately $1 - \alpha$. Therefore, the outcome of this random interval is an approximate $1 - \alpha$ confidence interval for β_i.

B. HYPOTHESIS TESTING

By using $\hat{\underline{\beta}}$, $\hat{\underline{\Sigma}}$ and the asymptotic normality of $\hat{\underline{\beta}}$, hypothesis tests for the parameters $P(\underline{x})$ and β_i $(i = 0,1,\ldots,k)$ may be easily constructed. As previously stated, asymptotically,

$$\underline{x}'\hat{\underline{\beta}} \sim N(\underline{x}'\beta,\ q^2(\underline{x})).$$

It follows that $(\underline{x}'\hat{\underline{\beta}} - c)/q(\underline{x})$ can be used as an asymptotic test statistic for testing H_0: $\underline{x}'\underline{\beta} \geq c$ for instance. However, H_0: $\underline{x}'\underline{\beta} \geq c$ is equivalent to H_0: $P(\underline{x}) \geq p_0$, where $c = \ln(p_0/(1 - p_0))$. Thus

$$[\underline{x}'\hat{\underline{\beta}} - \ln(p_0/(1 - p_0))]/q(\underline{x}) \qquad (3)$$

is the asymptotic test statistic for testing hypotheses relating $P(\underline{x})$ to some hypothesized value, p_0. Because of the asymptotic normality of $\underline{x}'\hat{\underline{\beta}}$, the test statistic for $P(\underline{x})$ is asymptotically normal with mean zero and variance one under the null hypothesis, i.e., when $P(x) = p_0$.

The statistical structure of hypothesis testing can be used to control the probability of making a wrong assumption concerning the relationship of $P(\underline{x})$ to p_0. Consideration will be given to the following two hypothesis testing formats, using a significance level of α:

(a) H_0: $P(\underline{x}) \leq p_0$
H_1: $P(\underline{x}) > p_0$

(b) H_0: $P(\underline{x}) \geq p_0$
H_1: $P(\underline{x}) < p_0$

No exact evaluation of the true value of $P(\underline{x})$ can be made. However it is important to know how hypothesis testing can be used most effectively in drawing safe conclusions about $P(\underline{x})$ in relation to p_0. First consider the hypothesis testing format in (a):

$$H_0: \; P(\underline{x}) \leq p_0 \quad \text{(null hypothesis)}$$
$$H_1: \; P(\underline{x}) > p_0 \quad \text{(alternative hypothesis)}$$

Rejection of H_0 when H_0 is true is called a Type I error. The probability of making such an error is α. Failing to reject H_0 when H_1 is true is a Type II error. The probability of making this error depends upon the sample size, values in the alternative hypothesis, and α. In this testing format a Type II error could be a dangerous error, causing the use of unsafe combinations of the independent variables, i.e., unsafe acceleration profiles.

Now consider the hypothesis testing format in (b):

$$H_0: \; P(\underline{x}) \geq p_0 \quad \text{(null hypothesis)}$$
$$H_1: \; P(\underline{x}) < p_0 \quad \text{(alternative hypothesis)}$$

Here the Type I and Type II errors are reversed from that in format (a). The Type I error is now the dangerous one. The structure of hypothesis testing is such that one simply specifies the probability of Type I error, whereas the probability of Type II error may be difficult to control. Thus, the hypothesis testing format of (b) is the safer one to use in making inferences concerning the size of $P(\underline{x})$ in relation to some specified p_0.

In some cases it might be desired to test whether a given parameter β_i is useful in the model, i.e., is nonzero. The hypothesis test for β_i

depends upon the fact that, asymptotically,

$$\hat{\beta}_i \sim N(\beta_i, v_i^2).$$

It follows that the appropriate test statistic is

$$(\hat{\beta}_i - b)/v_i \qquad (4)$$

for hypotheses relating β_i to some value b. The test statistic is asymptotically normal with mean zero and variance one under the null hypothesis, i.e., when $\beta_i = b$.

Another valuable hypothesis test is based on a likelihood ratio statistic which may be used for testing a reduced model, i.e., for testing whether or not a set of p certain parameters ($p \leq k$) are zero. This test was used in a recent study based on $-G_x$ acceleration data from subhuman primates [3]. The framework for this test involves the following two hypotheses:

H_0: $\beta_{i_1} = \beta_{i_2} = \ldots = \beta_{i_p} = 0$ (Reduced model)

H_1: β_i's unspecified (Full model)

The appropriate test statistic is

$$\lambda = L_1 - L_2 \qquad (5)$$

where

L_1 is -2 times the logistic likelihood using the maximum likelihood estimate obtained under H_0 and

L_2 is -2 times the logistic likelihood using the maximum likelihood estimate obtained under H_1.

The null hypothesis H_0 is rejected at a significance level α if $\lambda > \chi^2(\alpha, p)$, the upper α point of a chi-square distribution with p degrees of freedom.

IV. CRITICAL ENVELOPES

An inference of much importance is the assessment of $P(\underline{x})$, the true probability of injury for a given set of conditions $\underline{x}$, in relation to some small probability p_0. Such inferences can be made by the prediction of critical envelopes and by testing hypotheses concerning $P(\underline{x})$. A critical envelope can be defined as the set of all combinations of independent variables for which the predicted probability of injury is less than some given amount. However, variability in the predicted probability causes error variability in the prediction of the critical envelope. To predict safer critical envelopes such variability must be taken into account.

If $\hat{P}(\underline{x})$ is the predicted probability of injury, then $\hat{P}(\underline{x}) \leq p_0$ is equivalent to

$$x'\hat{\underline{\beta}} \leq \ln[p_0/(1 - p_0)]. \qquad (6)$$

Thus, choosing the combination of independent variables $\underline{x}$ such that (6) is true is equivalent to choosing $\underline{x}$ such that $\hat{P}(x) \leq p_0$. Let

$$C(p_0) = \{\underline{x}: \underline{x}'\hat{\underline{\beta}} = \ln[p_0/(1 - p_0)]\}.$$

Then $C(p_0)$ is the critical envelope corresponding to p_0. Such an envelope, however, does not adjust for the variability of $\hat{\underline{\beta}}$. This may result in the use of acceleration conditions that are unsafe.

A more conservative prediction of a critical envelope that leads to the use of safer conditions is

$$C_\alpha(p_0) = \{\underline{x}: \ \underline{x}'\hat{\underline{\beta}} + z_\alpha q(\underline{x}) = \ln[p_0/(1 - p_0)]\}. \qquad (7)$$

Utilization of combinations of the independent variables such that $\underline{x}$ is in the critical envelope implies that

$$\{1 + \exp(-[\underline{x}'\hat{\underline{\beta}} + z_\alpha q(\underline{x})])\}^{-1} = p_0. \qquad (8)$$

The left-hand side of equation (8) is the end point of an approximate $1 - \alpha$ right-sided confidence interval for the true $P(\underline{x})$. Thus one can be about $1 - \alpha$ confident that for any $\underline{x}$ in $C_\alpha(p_0)$,

$$P(\underline{x}) \leq \{1 + \exp[-(x'\hat{\underline{\beta}} + z_\alpha q(\underline{x}))]\}^{-1} = p_0.$$

By decreasing α, confidence may be increased by predicting a more conservative critical envelope. The additional safety achieved by such a critical envelope depends on the fact that an interval estimate, rather than a point estimate, is employed. The confidence interval used adjusts for the sampling variability of the estimated parameters in the injury model. It must be noted, however, that the smaller α is made, the more severe will be the constraints on the permitted acceleration profiles. In the extreme, selecting α extremely close to zero will result in no acceleration permitted. It is clear, therefore, that important tradeoffs will have to be made in establishing any critical envelopes to be adopted in practice.

V. EXAMPLES

This section presents examples of the inference results derived in the previous section. The data used in those examples is taken from the simulation study [4] previously mentioned. In each example of inference, results for Models A and B are given for sample sizes of N = 100 and N = 1000. Whenever possible, the examples list true parameter values along side the statistical results so that the accuracy of the inference procedures can be discerned.

Figure 5 provides an example of 95% confidence intervals for the probability $P(\underline{x})$ for a specific $\underline{x}$ vector. This vector was chosen because it corresponds to low probabilities (which tend to be of main interest) for both Models A and B. Figures 6 through 9 present, for both models and sample sizes, scatter plots of the length of various confidence intervals for $P(\underline{x})$ versus the corresponding true value of $P(\underline{x})$ for varying $\underline{x}$ vectors. It can be seen from these scatter plots that the length of the confidence interval for $P(\underline{x})$ tends to get smaller as $P(\underline{x})$ gets closer to one or zero. This is fortunate since concern centers on small probabilities of injury.

Figure 10 lists 95% confidence intervals for parameter β_1 of Models A and B. Figures 11 and 12 exhibit the results of hypothesis tests of $P(\underline{x})$ and β_1, respectively. The specific $\underline{x}$ vector used for the confidence intervals in Figure 5 was also used for the hypothesis test in Figure 11 because it corresponds to low probabilities.

Figure 13, using Model B with a sample size of 100, illustrates the test of a reduced model defined by setting the three smallest model parameters to zero (i.e., $\beta_1 = \beta_2 = \beta_3 = 0$). The three parameters are

Model	Sample Size	True Value	Estimated Value	Confidence Interval	Confidence Interval Length
A	100	0.1104	0.0578	[0.0147, 0.2027]	0.1880
A	1000	0.1104	0.0970	[0.0667, 0.1391]	0.0724
B	100	0.0165	0.0127	[0.0015, 0.0964]	0.0949
B	1000	0.0165	0.0123	[0.0066, 0.0227]	0.0161

Figure 5: 95% Confidence Intervals for $P(\underline{x})$ with $\underline{x}$ = (1, -0.0862, -0.2621, 0.9114, -0.7660, -0.1363, -0.4655)

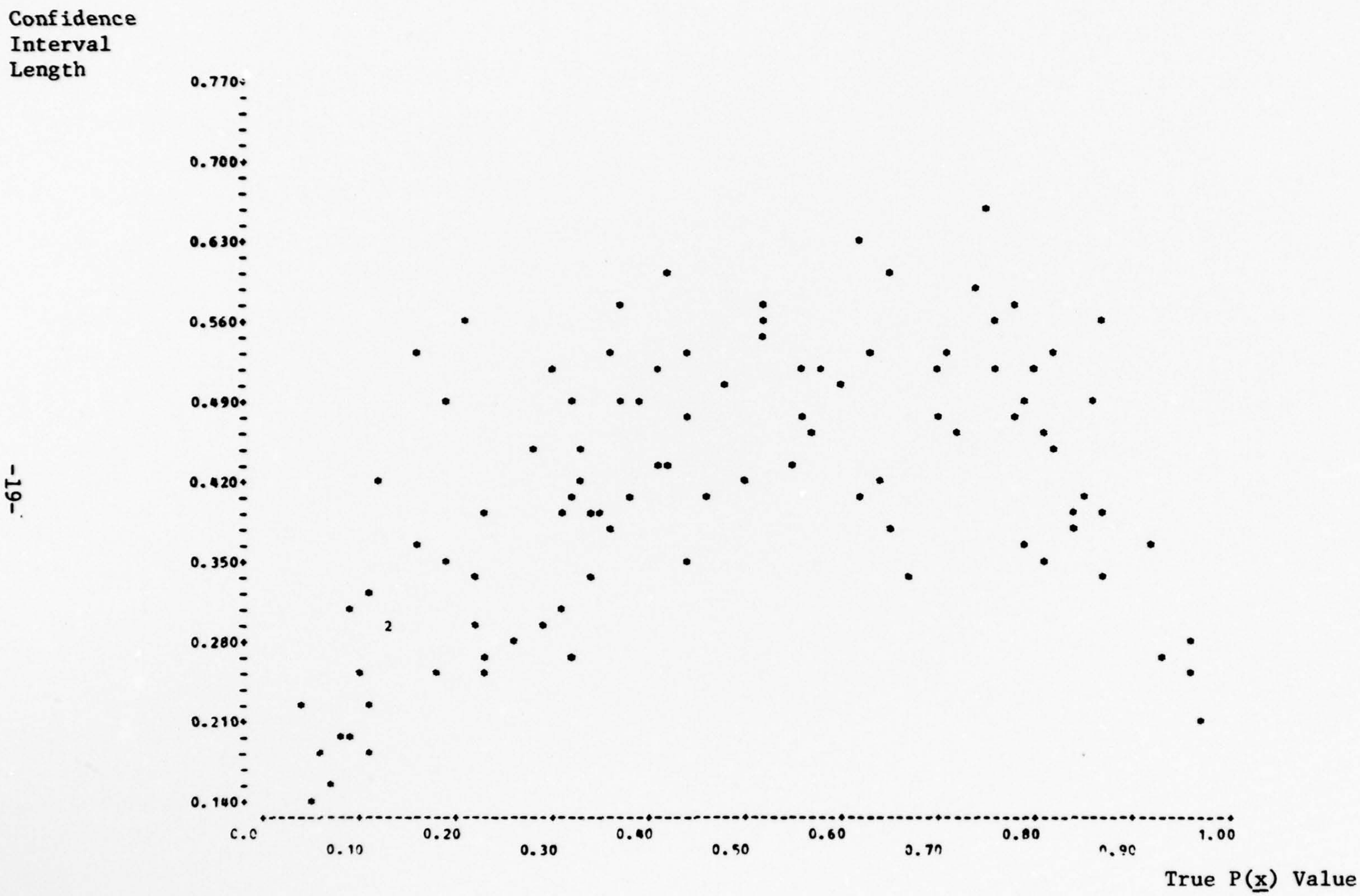


Figure 6: Confidence Interval Length Versus True P(x) Value for Model A and Sample Size of 100

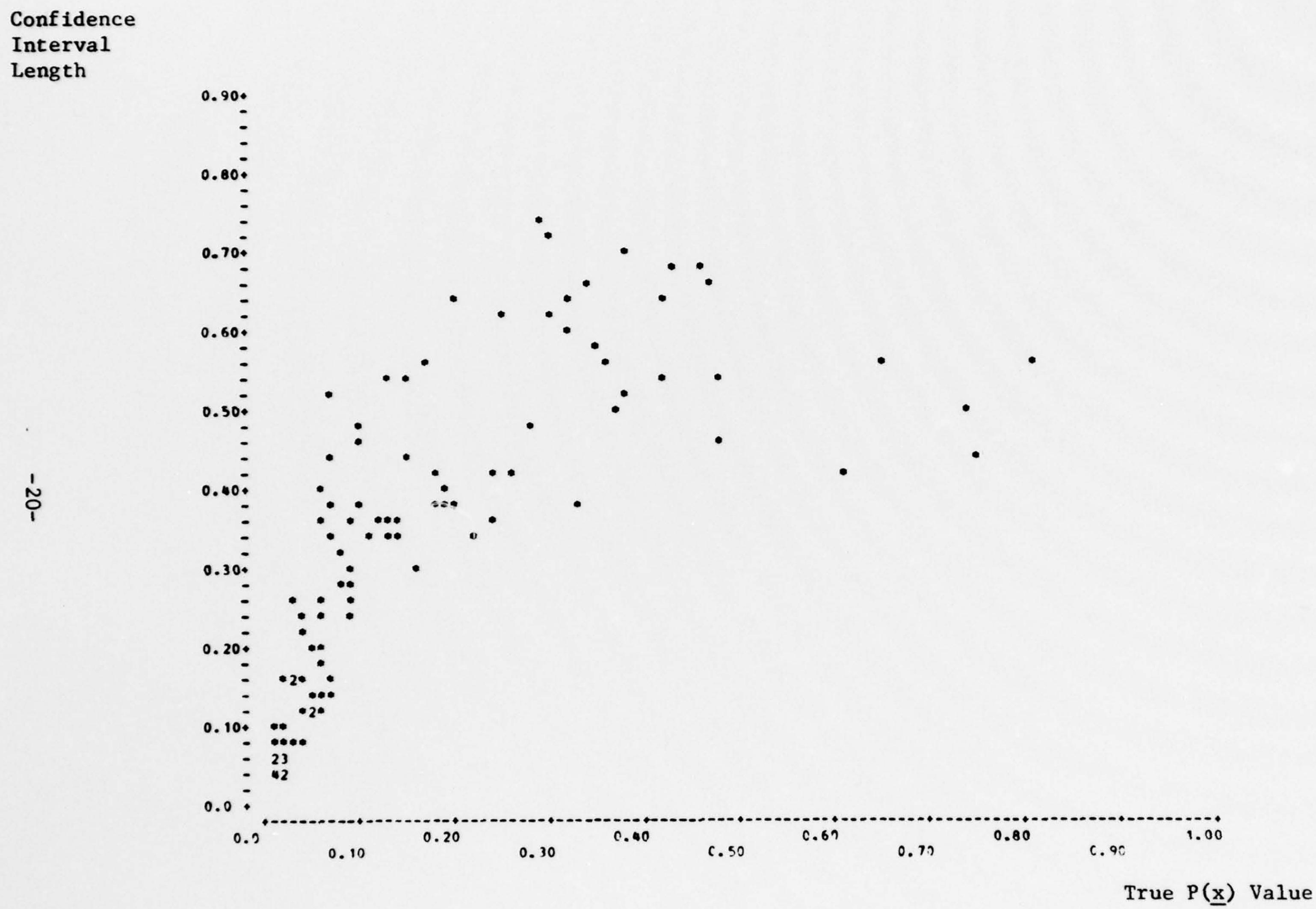


Figure 7: Confidence Interval Length Versus True P(x) Value for Model B and Sample Size of 100

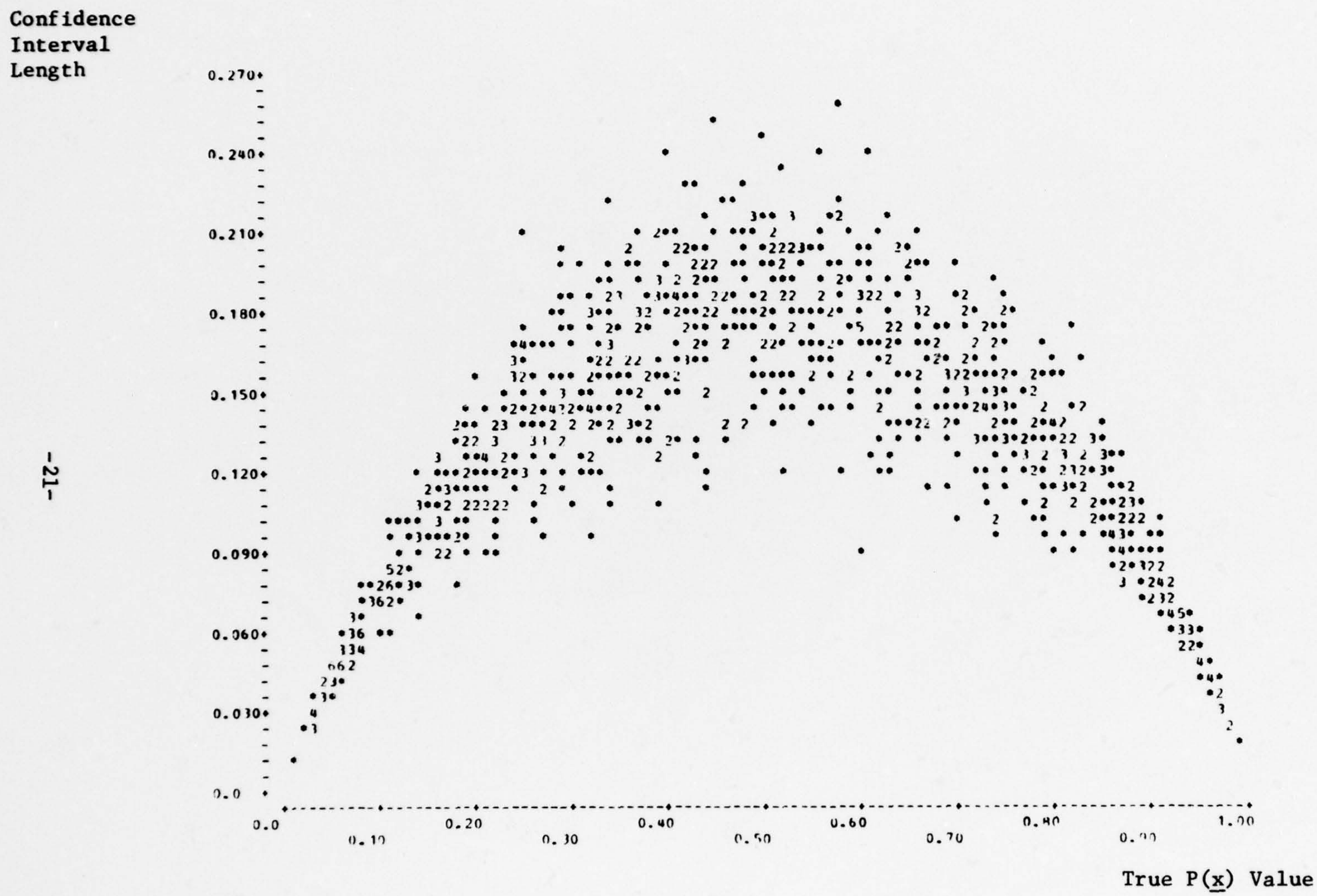


Figure 8: Confidence Interval Length Versus True P(x) Value for Model A and Sample Size of 1000

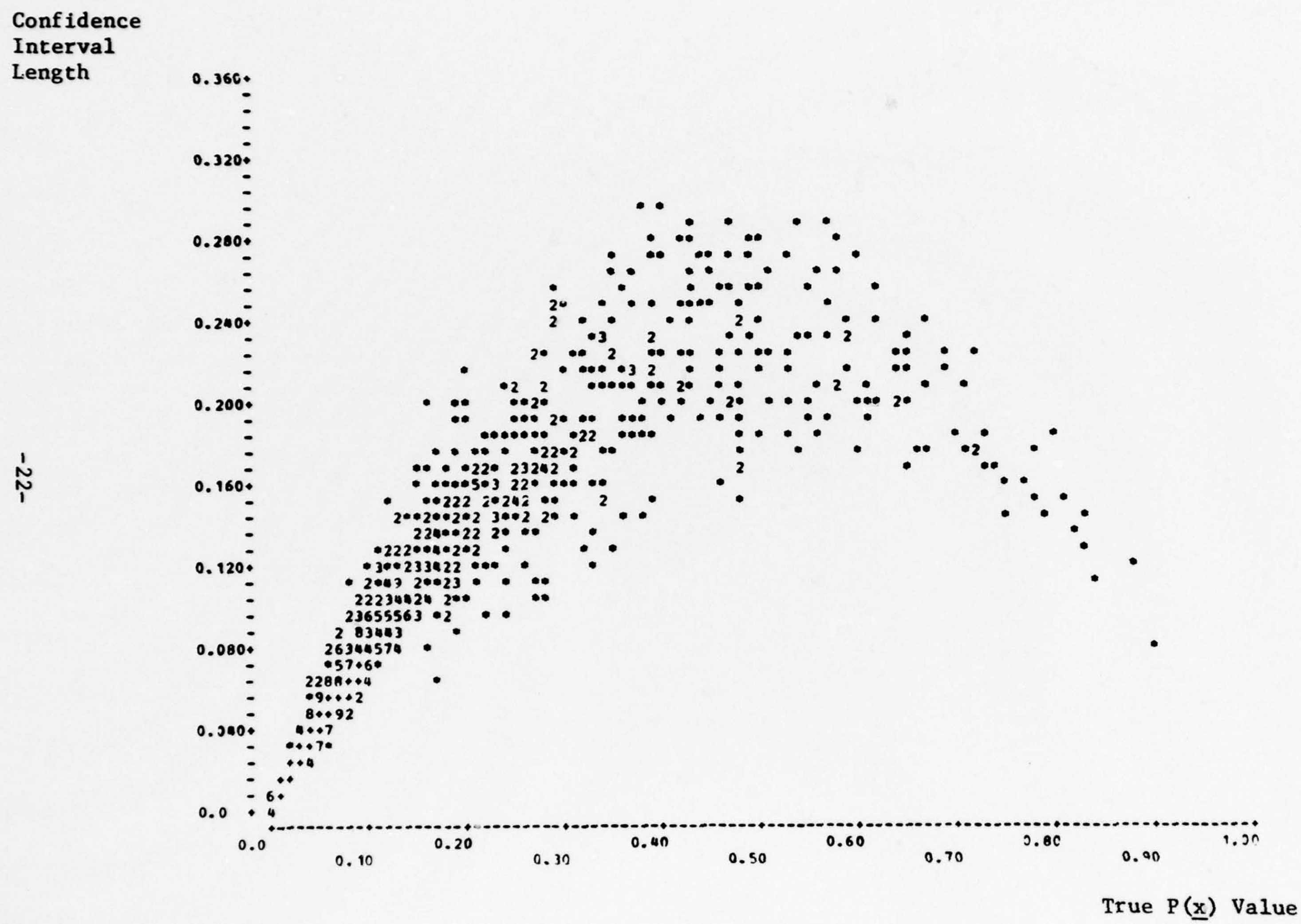


Figure 9: Confidence Interval Length Versus True P(x) Value for Model B and Sample Size of 1000

Model	Sample Size	True Value	Estimated Value	Confidence Interval	Confidence Interval Length
A	100	-0.25	-0.4270	[-1.3035, 0.4495]	1.7530
A	1000	-0.25	-0.1948	[-0.2748, 0.0605]	0.3353
B	100	-0.25	-0.0563	[-1.1845, 1.0719]	2.2564
B	1000	-0.25	-0.2609	[-0.5886, 0.0667]	0.6553

Figure 10: 95% Confidence Intervals for β_1

Model	Sample Size	True Probability	Estimated Probability	Test Statistic	p-value	Decision
A	100	0.1104	0.0578	0.2155	0.5850	fail to reject H_0
A	1000	0.1104	0.0970	3.4264	0.9997	fail to reject H_0
B	100	0.0165	0.0127	-1.3068	0.0951	fail to reject H_0
B	1000	0.0165	0.0123	-4.5380	0.0000^+	reject H_0

Figure 11: Hypothesis Tests of H_0: $P(\underline{x}) \geq 0.05$ versus H_1: $P(\underline{x}) < 0.05$ ($\alpha = 0.05$)

Model	Sample Size	True Value	Estimated Value	Value of Test Statistic	p-value	Decision
A	100	-0.25	-0.4270	-0.9532	0.3400	fail to reject H_0
A	1000	-0.25	-0.1948	-1.4954	0.1350	fail to reject H_0
B	100	-0.25	-0.0563	-0.0978	0.4610	fail to reject H_0
B	1000	-0.25	-0.2609	-1.5610	0.0595	fail to reject H_0

Figure 12: Hypothesis Tests of H_0: $\beta_1 = 0$ versus H_1: $\beta_1 \neq 0$ ($\alpha = 0.05$)

H_0: $\beta_1 = \beta_2 = \beta_3 = 0$ (Reduced Model)

H_1: β_i's unspecified (Full Model)

The test statistic is $\lambda = L_1 - L_2$

where L_1 = -2 times the logistic likelihood using the maximum likelihood estimate, $\underline{\hat{\hat{\beta}}}$, obtained under H_0

$= 66.86$

and L_2 = -2 times the logistic likelihood using the maximum likelihood estimate, $\underline{\hat{\beta}}$, obtained under H_1

$= 65.60$

Thus,

$$\lambda = 1.26 < \chi^2(.05, 3) = 7.81$$

so the null hypothesis,

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

cannot be rejected at significance level $\alpha = 0.05$.

Figure 13: Test of a Reduced Model

shown to be not statistically significant at the 0.05 significance level.

Since the actual critical envelopes for the simulated data are six-dimensional, it is impossible to exhibit them graphically. However, some information can be provided by the two-dimensional plots shown in Figures 14 and 15. Since x_5 and x_6 are the two strongest predictor variables, they were allowed to vary to form plots, while x_1, x_2, x_3, and x_4 were held fixed at the values used in the previous inferences concerning $P(\underline{x})$. The two-dimensional projection of the critical envelope (a straight line) derived in [3] is plotted on the same graph as the corresponding two-dimensional safer critical envelope (a curved line) derived in this technical report. Figure 14 illustrates the two-dimensional critical envelope plots corresponding to Model B with $\alpha = 0.05$, $p_0 = 0.10$, and a sample size of 100 observations. Figure 15 illustrates similar plots except that the sample size was increased to 1000 observations. Note that as the sample size increases from 100 to 1000 observations, the two different types of critical envelope become closer. The main factors influencing the shape of the critical envelopes derived in this report are:

(1) the confidence level of the corresponding one-sided confidence interval, α,

(2) the critical envelope probability, p_0,

(3) the statistical significance of the two plotted predictor variables,

(4) the structure of the variance-covariance matrix of $\hat{\underline{\beta}}$,

and (5) the sample size used to estimate $\underline{\beta}$.

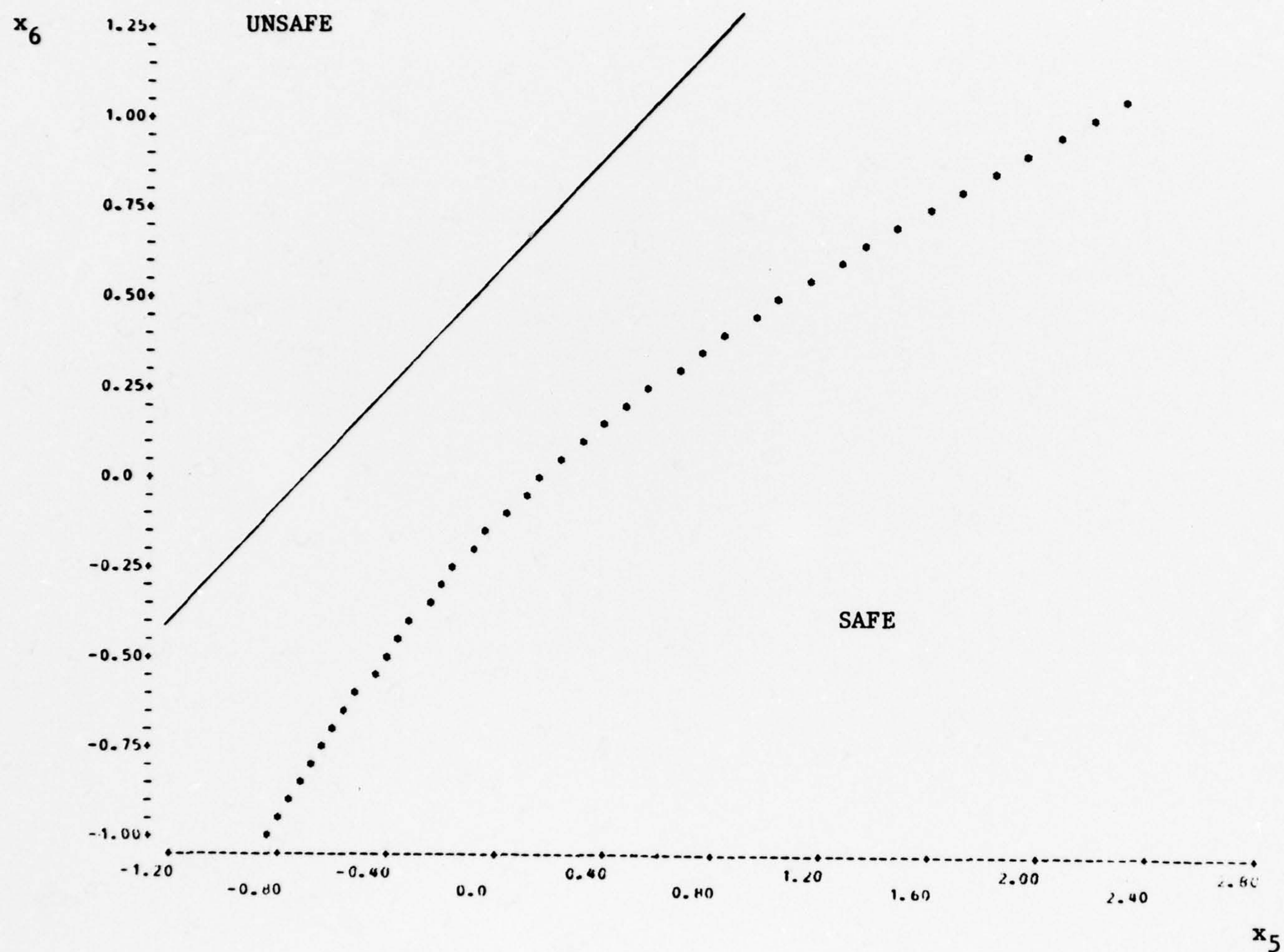


Figure 14: Two-dimensional Projection of Critical Envelopes for Model B and a Sample Size of 100 ($\alpha = 0.05$, $p_0 = 0.10$)

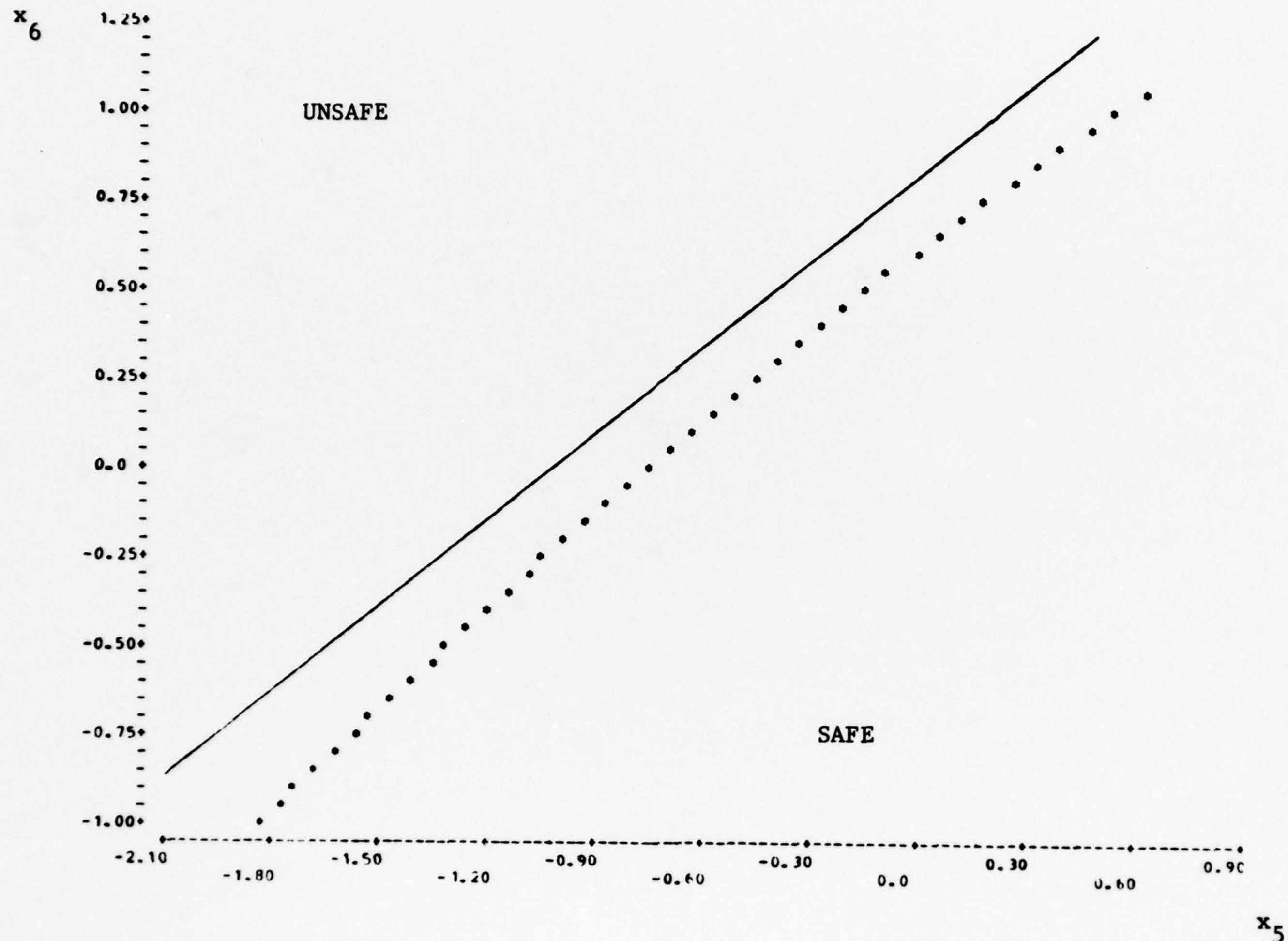


Figure 15: Two-dimensional Projection of Critical Envelopes for Model B and a Sample Size of 1000 ($\alpha = 0.05$, $p_0 = 0.10$)

VI. DISCUSSION

The statistical inference procedures discussed in this technical report account for variability in the data. Because these procedures are asymptotic, they are more accurate for large samples than for small samples. Nonetheless, the use of these inference procedures in computing injury threshold levels is far safer than the use of point estimates, which do not account for sampling error.

The inferences described in this report were formulated through the use of maximum likelihood theory. Such statistical inferences utilize only sample information. However, it may be desirable to combine a researcher's prior knowledge about an experiment with the experimental sample information by means of Bayesian analysis. The assessment of a prior multivariate probability distribution for vector $\underline{\beta}$ could be very difficult. However, for a given $\underline{x}$ vector of concern, an experimenter's prior knowledge of $P(\underline{x})$ should be relatively easy to quantify. Once a prior distribution of values for $P(\underline{x})$ is obtained, the prior information can be combined with the sample information to produce inferences that reflect both prior and sample information.

VII. REFERENCES

[1] Jones, R. H., "Probability Estimation Using a Multinomial Logistic Function", J. Statist. Comput. Simul., Vol. 3, pp. 315-329 (1975).

[2] Smith, D. E., "Research on Construction of a Statistical Model for Predicting Impact Acceleration Injury", Technical Report No. 102-2, Desmatics, Inc., 1976.

[3] Smith, D. E., "An Examination of Statistical Impact Acceleration Injury Prediction Models Based on $-G_x$ Accelerator Data from Subhuman Primates", Technical Report No. 102-6, Desmatics, Inc., 1978.

[4] Smith, D. E. and Gardner, R. L., "A Study of Estimation Accuracy When Using a Logistic Model for Prediction of Impact Acceleration Injury", Technical Report No. 102-5, Desmatics, Inc., 1978.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 102-7	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) STATISTICAL INFERENCE PROCEDURES FOR A LOGISTIC IMPACT ACCELERATION INJURY PREDICTION MODEL		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) John J. Peterson and Dennis E. Smith		8. CONTRACT OR GRANT NUMBER(s) N00014-74-C-0154
9. PERFORMING ORGANIZATION NAME AND ADDRESS Desmatics, Inc. P. O. Box 618 State College, PA 16801		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 207-037
11. CONTROLLING OFFICE NAME AND ADDRESS Biophysics Program (Code 444) Office of Naval Research Arlington, VA 22217		12. REPORT DATE December 1978 13. NUMBER OF PAGES 33
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Distribution of this report is unlimited.

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Impact Acceleration Injury
Logistic Prediction Model
Statistical Inference

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

This report describes statistical inference procedures relating to a logistic impact acceleration injury prediction model. The inference procedures discussed are interval estimation, hypothesis testing, and critical envelope estimation. Fundamentally, a critical envelope defines the independent variable combination such that the probability of injury is unlikely to exceed a prespecified value. Examples of the inference procedures discussed are given using data from a Monte Carlo simulation study.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)