

AD-A095 299

DESMATICS INC STATE COLLEGE PA

F/G 12/1

COMPUTATIONAL ASPECTS OF INCORPORATING AUXILIARY INFORMATION IN--ETC(U)

FEB 81 J J PETERSON, D E SMITH

N00014-79-C-0128

UNCLASSIFIED

TR-112-7

NL

1 of 1
ALL
AC095299



END
DATE
FILMED
3 81
DTIC

AD A09529



DESMATICS, INC. ✓

P. O. Box 618

State College, Pa. 16801

Phone: (814) 238-9621

Applied Research in Statistics - Mathematics - Operations Research

6
COMPUTATIONAL ASPECTS OF INCORPORATING
AUXILIARY INFORMATION INTO AN IMPACT
ACCELERATION INJURY PREDICTION MODEL.

by

10 John J. Peterson

Dennis E. Smith

14 7K-11-7
9 TECHNICAL REPORT NO. H12-7

11 February 1981

This study was supported by the Office of Naval Research
under Contract No. N00014-79-C-0128 Task No. NR 207-037

15
Reproduction in whole or in part is permitted
for any purpose of the United States Government

Approved for public release; distribution unlimited

505

TABLE OF CONTENTS

	<u>Page</u>
I. INTRODUCTION	1
II. COMPUTATIONAL PROCEDURES	2
A. PRIOR INFORMATION	2
B. PARAMETER CONSTRAINTS	5
III. A POSSIBLE EXPERIMENT	8
A. SOME ASSUMPTIONS	8
B. MODEL REPRESENTATION	10
C. CORRELATION	12
IV. DISCUSSION	14
V. REFERENCES	15
APPENDIX	16

Accession For

1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	23
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	----

I. INTRODUCTION

A previous report [3] presented a general outline of procedures for simultaneously incorporating various sources of auxiliary information into an impact acceleration injury prediction model. This report, which serves as a companion volume, discusses computational aspects of the procedures. Specifically, the report presents computational procedures which allow the application of commonly used nonlinear estimation programs found in statistical packages such as BMDP [1] and SAS [2].

Throughout this report, reference will be made to equations in the previous report [3]. These equations, identified as (1) through (17) in that report, will be referenced herein by the same numbers as a matter of convenience. Equations in this report will therefore begin with reference number (18). In addition, the notation used in the previous report will be adopted without repeating its definition. Therefore, the reader may find it convenient (or perhaps necessary) to have that report readily available.

II. COMPUTATIONAL PROCEDURES

To estimate the model in (17) by using a statistical package such as SAS or BMDP, one should iteratively compute, for each observation:

(i) the model derivatives of (17) with respect to the parameters $\underline{\beta}_1$ and ρ

and (ii) the regression weights, $[p_i(1 - p_i)]^{-1}$

The derivatives of (17) with respect to $\underline{\beta}_1$ and ρ are

$$\frac{\partial}{\partial \underline{\beta}_1} p_i = \phi(U_i(\underline{\beta}_1, \rho))(1 - \rho^2)^{-\frac{1}{2}} \underline{x}_i$$

$$\frac{\partial}{\partial \rho} p_i = \phi(U_i(\underline{\beta}_1, \rho)) \cdot \{-(1 - \rho^2)^{-\frac{1}{2}} s_i + [\rho/(1 - \rho^2)]U_i(\underline{\beta}_1, \rho)\}$$

where $\phi(U) = \frac{d}{dU} \Phi(U)$.

This paragraph concludes the discussion of the problem without regard to prior information.

A. PRIOR INFORMATION

The model parameters in (6) or (7), as the case may be, can be estimated conveniently and efficiently by using the model in (17) to incorporate the prior information into the parameter estimates. Here again the intuition behind this approach comes from the likelihood of the observations. Consider the likelihood in (12). The form of this likelihood suggests the models:

$$y_{1i} = \Phi[(1 - \rho^2)^{-1/2}(\underline{x}_i' \underline{\beta}_1 - \rho s_i)] + \xi_i \quad (18a)$$

$$y_{2i} = \underline{x}_i' \underline{\beta}_2 + \varepsilon_{2i} \quad (18b)$$

where $E(\xi_i) = 0$, $\text{Var}(\xi_i) = p_i(1 - p_i)$, $p_i = \Phi[U_i(\underline{\beta}_1, \rho)]$

$$E(\varepsilon_{2i}) = 0, \text{Var}(\varepsilon_{2i}) = \sigma_2^2.$$

Note that the ξ_i and ε_{2i} error terms are stochastically independent, since y_{1i}/s_i and y_{2i} are independent. This fact makes the analysis by statistical packages convenient since their nonlinear regression algorithms cannot analyze correlated observations directly.

As stated in the previous report [3], it is assumed that the prior information can be in the form of

- (i) a priori estimates of some function of the model parameters
- and/or (ii) a priori knowledge in the form of model parameter equality constraints.

The technique used to model the prior information discussed in the previous report can be applied to the models in (18a) and (18b) to form an overall model of direct and auxiliary information.

Let the prior estimates of some function of the model parameters be expressed by (6), where

$$\underline{\alpha} = \begin{pmatrix} \underline{\beta} \\ \rho \end{pmatrix}.$$

Recall that it is assumed that the prior information embodied in $\underline{r}$ is stochastically independent of

$$\begin{pmatrix} y_{11} \\ y_{21} \end{pmatrix}, \dots, \begin{pmatrix} y_{1n} \\ y_{2n} \end{pmatrix}.$$

$$\text{Let } \underline{f}_i(\underline{\beta}, \rho) = \begin{bmatrix} f_{1i}(\underline{\beta}, \rho) \\ f_{2i}(\underline{\beta}, \rho) \end{bmatrix} = \begin{bmatrix} \phi(U_i(\underline{\beta}_1, \rho)) \\ \underline{x}_i' \underline{\beta}_2 \end{bmatrix}, \underline{f}(\underline{\beta}, \rho) = \begin{bmatrix} \underline{f}_1(\underline{\beta}, \rho) \\ \vdots \\ \underline{f}_n(\underline{\beta}, \rho) \end{bmatrix}$$

$$\underline{\eta}_1 = \begin{bmatrix} \xi_1 \\ \epsilon_{21} \end{bmatrix}, \quad \underline{\eta} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}.$$

Then all the information, barring parameter constraints, can be modeled as

$$\underline{r} = \underline{g}(\underline{\beta}, \rho) + \underline{v} \quad (19a)$$

$$\underline{y} = \underline{f}(\underline{\beta}, \rho) + \underline{\eta} \quad (19b)$$

where $E(\underline{v}) = 0$, $\text{Var}(\underline{v}) = \underline{\Psi}$

$$\text{and } E(\underline{\eta}) = 0, \text{Var}(\underline{\eta}) = \underline{V}^* = \begin{bmatrix} \underline{V}_1^* & 0 \\ 0 & \underline{V}_n^* \end{bmatrix}, \quad \underline{V}_1^* = \begin{bmatrix} p_1(1 - p_1) & 0 \\ 0 & \sigma_2^2 \end{bmatrix},$$

$$p_1 = \phi(U_1(\underline{\beta}_1, \rho)).$$

The vector $\underline{r}$ is conceptually and computationally considered as an observation. Due to the structure of most statistical packages, this concept can be computationally exploited only if $\underline{\Psi}$ is a diagonal matrix, where the diagonal elements of $\underline{\Psi}$ are considered as weights for the observations in $\underline{r}$. Even if $\underline{\Psi}$ is a diagonal matrix, it is probably the easiest for the user if he or she transforms the elements of $\underline{r}$ so that the transformed elements have the identity matrix as covariance matrix. This is done in this report.

The elements of $\underline{r}$ can be transformed by premultiplying $\underline{r}$ by a matrix $\underline{\Gamma}$ such that

$$\underline{\Gamma}' \underline{\Gamma} = \underline{\Psi}^{-1}.$$

If $\underline{r}^*$ denotes the transformed elements of $\underline{r}$, then (19a) can be reexpressed as

$$\underline{r}^* = \underline{\Gamma} \underline{r} = \underline{g}^*(\underline{\beta}, \rho) + \underline{\Gamma} \underline{v}$$

where

$$\underline{g}^*(\underline{\beta}, \rho) = \underline{\Gamma} \underline{g}(\underline{\beta}, \rho)$$

and

$$E(\underline{\Gamma} \underline{v}) = 0, \text{Var}(\underline{\Gamma} \underline{v}) = \underline{\Gamma}' (\underline{\Gamma}' \underline{\Gamma})^{-1} \underline{\Gamma}' = \underline{I}.$$

B. PARAMETER CONSTRAINTS

To include constraints on the parameters, let

$$(\underline{\beta}, \rho) = \underline{h}(\underline{\theta}) \tag{20}$$

where

$$\dim \underline{\theta} \leq \dim(\underline{\beta}, \rho).$$

Then, it follows that, defining $\underline{v}^* = \Gamma \underline{v}$,

$$\underline{r}^* = \underline{g}^*(\underline{h}(\underline{\theta})) + \underline{v}^* \quad (21a)$$

$$\underline{y} = \underline{f}(\underline{h}(\underline{\theta})) + \underline{\eta} \quad (21b)$$

Note that indicator variables can be used to facilitate the computation of the estimates of $(\underline{\beta}, \rho)$ with a statistical package.

For example, consider the case without constraints and let

$$z_i = \delta_{1i} y_{1i} + \delta_{2i} y_{2i} + (1 - \delta_{1i})(1 - \delta_{2i}) r_j^* \quad (22)$$

$$k_i(\underline{\beta}, \rho) = \delta_{1i} f_{1i}(\underline{\beta}, \rho) + \delta_{2i} f_{2i}(\underline{\beta}, \rho) + (1 - \delta_{1i})(1 - \delta_{2i}) g_j^*(\underline{\beta}, \rho)$$

$$e_i = \delta_{1i} \xi_i + \delta_{2i} \epsilon_{1i} + (1 - \delta_{1i})(1 - \delta_{2i}) v_j^*$$

$$\text{where } \delta_{1i} = \begin{cases} 1 & \text{if } z_i = y_{1i}, k_i(\underline{\beta}, \rho) = f_{1i}(\underline{\beta}, \rho), \text{ and } e_i = \xi_i \\ 0 & \text{otherwise} \end{cases}$$

$$\delta_{2i} = \begin{cases} 1 & \text{if } z_i = y_{2i}, k_i(\underline{\beta}, \rho) = f_{2i}(\underline{\beta}, \rho), \text{ and } e_i = \epsilon_{2i} \\ 0 & \text{otherwise} \end{cases}$$

and $i = 1, \dots, n$, $j = 1, \dots, m$, $m = \dim \underline{r}^*$.

Here r_j , g_j^* , and v_j are the j^{th} elements of $\underline{r}$, $\underline{g}^*$, and $\underline{v}$ respectively.

Now rewrite (22) as

$$\underline{z} = \underline{k}(\underline{\beta}, \rho) + \underline{e} \quad (23)$$

$$\text{where } \underline{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \underline{k}(\underline{\beta}, \rho) = \begin{bmatrix} k_1(\underline{\beta}, \rho) \\ \vdots \\ k_n(\underline{\beta}, \rho) \end{bmatrix}, \underline{e} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix}$$

and $E(\underline{e}) = 0$, $\text{Var}(e_i) = \delta_{1i} p_i (1 - p_i) + \delta_{2i} \sigma_2^2 + (1 - \delta_{1i})(1 - \delta_{2i})$,
 with $e_1, \dots, e_n$ independent.

Next, if parameter constraints exist as in (20), just substitute $\underline{h}(\underline{\theta})$ for $(\underline{\beta}, \rho)$ in the model in (23). The above indicator variables will enable the statistical package to directly analyze all the information modeled in (23) via standard nonlinear, least squares algorithms. One merely needs to iteratively compute:

(i) model derivatives of (23) with respect to $(\underline{\beta}, \rho)$ or

$\underline{\theta}$ if constraints are to be used,

and (ii) the regression weights $[\text{Var}(e_i)]^{-1}$.

III. A POSSIBLE EXPERIMENT

To elucidate the application of estimating the model in (23), an example of a possible impact acceleration injury experiment with n observations may be considered. Suppose that for $i = 1, \dots, n$:

(i) $\underline{x}_i = (1, x_{1i})'$ where x_{1i} denotes peak sled acceleration for the i^{th} subject,

(ii) y_{1i} is a dichotomous injury observation for i^{th} subject, with value of 1 for injury, 0 for no injury,

and (iii) y_{2i} is a continuous preinjury measurement for the i^{th} subject.

Then the model of the empirical data as given in (19b) has

$$\underline{\beta} = \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \end{bmatrix}, \quad \underline{\beta}_1 = \begin{bmatrix} \beta_{01} \\ \beta_{11} \end{bmatrix}, \quad \underline{\beta}_2 = \begin{bmatrix} \beta_{02} \\ \beta_{12} \end{bmatrix}$$

A. SOME ASSUMPTIONS

Suppose further, that for a particular level of peak sled acceleration, x_0 say, a good prior unbiased estimate of injury probability, p_0 , is available. Let this prior injury probability and peak sled acceleration level be related by

$$p_0 = \phi(\beta_{01} + \beta_{11}x_0).$$

Assume that the variance of p_0 is known or estimated to be ψ .

Thus, the prior information as modeled in (19a) can be written as

$$\underline{r}^* = \underline{g}^*(\underline{\beta}, \rho) + \underline{\Gamma}v$$

where $\underline{\Gamma}$ is a 1×1 matrix estimated by ψ^{-1} , $\underline{r}^*$ is a 1×1 vector equal to $\psi^{-1}p_0$, and $\underline{g}^*(\underline{\beta}, \rho) = \psi^{-1}[\Phi([1, x_0, 0, 0]\underline{\beta})]$.

Finally, assume that the elements of $\underline{\beta}$ are constrained by the equation $\beta_{12} = \sigma_2 \beta_{11}$. This constraint can arise quite naturally from the following situation. Suppose that the event of a preinjury measurement, (e.g., change in evoked potential response) exceeding some critical, prespecified value, y_0 , is always at least as likely as the occurrence of an injury, for all levels of peak sled acceleration. This can be mathematically represented by

$$\Pr(y_{2i} > y_0 | \underline{x}) \geq \Pr(y_{1i} = 1 | \underline{x}) \text{ for all } \underline{x}.$$

Now,

$$\Pr(y_{2i} > y_0 | \underline{x}) = \Phi[-(y_0 - \underline{x}'\underline{\beta}_2)/\sigma_2]$$

$$\text{and } \Pr(y_{1i} = 1 | \underline{x}) = \Phi(\underline{x}'\underline{\beta}_1).$$

Therefore, for all $\underline{x}$,

$$\Phi[-(y_0 - \underline{x}'\underline{\beta}_2)/\sigma_2] \geq \Phi(\underline{x}'\underline{\beta}_1).$$

However, this can only be true if

$$\beta_{12}/\sigma_2 = \beta_{11}.$$

The constraint imposed by the vector valued function $\underline{h}$ can be described

by

$$\underline{\beta} = \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{02} \\ \beta_{12} \\ \rho \end{bmatrix} = \underline{h}(\underline{\theta}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{\theta},$$

where

$$\underline{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{bmatrix} = \begin{bmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{02} \\ \rho \end{bmatrix}.$$

However, $\underline{h}$ depends upon σ_2 , which is unknown. In other words, the prior information, in terms of the parameter constraints, is not complete. In this particular case, the best that can be done is to estimate $\underline{h}$ by $\hat{\underline{h}}$,

where

$$\hat{\underline{h}}(\underline{\theta}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \tilde{\sigma}_2 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \underline{\theta}$$

and $\tilde{\sigma}_2$ is the minimum variance, unbiased estimator of σ_2 based on the preinjury data, i.e. $\tilde{\sigma}_2 = \frac{n}{n-2} \hat{\sigma}_2$, $\hat{\sigma}_2$ being the maximum likelihood estimate of σ_2 .

B. MODEL REPRESENTATION

The model representation for computer analysis is then given (for $i = 1, \dots, n$ and $j = 1$) by:

$$z_i = \delta_{1i} y_{1i} + \delta_{2i} y_{2i} + (1 - \delta_{1i})(1 - \delta_{2i}) r_j^*$$

$$k_i(\underline{\beta}, \rho) = \delta_{1i} f_{1i}(\hat{\underline{h}}(\underline{\theta})) + \delta_{2i} f_{2i}(\hat{\underline{h}}(\underline{\theta})) + (1 - \delta_{1i})(1 - \delta_{2i}) g_j^*(\hat{\underline{h}}(\underline{\theta}))$$

$$e_i = \delta_{1i} \xi_i + \delta_{2i} \epsilon_{2i} + (1 - \delta_{1i})(1 - \delta_{2i}) v_j^*$$

where δ_{11} and δ_{21} were defined previously, and

$$\begin{bmatrix} f_{11}(\hat{h}(\theta)) \\ f_{21}(\hat{h}(\theta)) \end{bmatrix} = \begin{bmatrix} \phi(u_1(\beta_1, \rho)) \\ \beta_{02} + x_{11}\tilde{\sigma}_2\beta_{11} \end{bmatrix},$$

$$r_j^* = \psi^{-1}p_0,$$

$$g_j^*(\hat{h}(\theta)) = \psi^{-1}\phi(\beta_{01} + x_0\beta_{11}).$$

If a separate estimate of β_1 based on the dichotomous injury data and separate estimates of β_2 and σ_2 based on the continuous preinjury data are available, then a "quick-and-dirty" estimate of ρ can be computed.

Note that

$$E(\epsilon_{21}y_{11}) = -\rho\sigma_2\phi(x_1'\beta_1) \quad (24)$$

$$\text{Var}(\epsilon_{21}y_{11}) = \sigma_2^2\{\phi(x_1'\beta_1) - \rho^2[\phi(x_1'\beta_1)\phi(x_1'\beta_1) + \phi^2(x_1'\beta_1)]\}$$

The results in (24) are derived in the appendix of this and the previous technical report [3]. From (24) it is easy to see that

$$E\{-\epsilon_{21}y_{11}(\sigma_2\phi(x_1'\beta_1))^{-1}\} = \rho.$$

Thus,

$$-\sum_{i=1}^n \epsilon_{2i}y_{1i} / [\phi(x_1'\tilde{\beta}_1)n\tilde{\sigma}_2^2] \quad (25)$$

is an approximately unbiased estimate of ρ . Here $\tilde{\beta}_1$ is the probit estimate of β_1 based on the dichotomous injury data only and $\tilde{\sigma}_2^2$ is the unbiased estimate of σ_2^2 ,

$$\frac{1}{(n-2)} \sum_{i=1}^n \epsilon_{2i}^2.$$

Unfortunately, the estimate in (25) can sometimes have absolute values exceeding one. A more refined (approximately unbiased) estimate of ρ can be easily obtained by using an iterative least squares statistical algorithm. To compute an iteratively reweighted, least squares estimate of ρ , simply let

$$\begin{bmatrix} \epsilon_{21} y_{11} \\ \vdots \\ \epsilon_{2n} y_{1n} \end{bmatrix} \quad (26)$$

be the dependent observations and let

$$\begin{bmatrix} \sim \sigma_2 \phi(\underline{x}_1' \tilde{\beta}_1) \\ \vdots \\ \sim \sigma_2 \phi(\underline{x}_{n-1}' \tilde{\beta}_1) \end{bmatrix} \quad (27)$$

be the independent observations. Compute the initial weights from the reciprocals of the variances of $\epsilon_{2i} y_{1i}$ by picking an initial value for ρ . The model derivatives are estimated by (27).

C. CORRELATION

Thus far, all models have assumed that ρ is constant for all $\underline{x}$. An approximate chi-square test can be conducted to test the null hypothesis that ρ does not depend upon $\underline{x}$. From (24) the means and variances of $\epsilon_{2i} y_{1i}$ can be computed. Let

$$\mu_1 = E(\epsilon_{21} y_{11}) ,$$

$$\sigma_1 = \text{Var}(\epsilon_{21} y_{11}) ,$$

$$w_1 = \frac{\epsilon_{21} y_{21} - \mu_1}{\sigma_1} ,$$

$$\text{and } \bar{w} = \frac{1}{n} \sum_{i=1}^n w_i .$$

Then $w_1, \dots, w_n$ are stochastically independent, each with mean zero and variance equal to 1. It follows that

$$\sum_{i=1}^n (w_i - \bar{w})^2$$

has an approximate chi-square distribution with $(n - 1)$ degrees of freedom. Of course all the μ_i 's and σ_i 's are not known, but they can be estimated by substituting the estimates of the model parameters via the equations in (24). Replacing the μ_i 's and σ_i 's by their estimates should still yield a reasonably powerful test against significant departures from the null hypothesis for moderate to large samples sizes.

If the chi-square test rejects the null hypothesis, then it is still possible to estimate a probit model where ρ depends upon $\underline{x}$. However in this case, it would be desirable to increase the sample size. This is due to the addition of the extra parameters used to model ρ . The computation of the model parameter estimates would also be considerably more difficult to obtain.

IV. DISCUSSION

The purpose of using auxiliary information is to improve the estimate of β_1 . In some estimation situations, however, the use of auxiliary information may result in estimates of β_1 that are slightly less accurate than if the auxiliary information were not used at all. If the constraint information is correct and the prior estimates of β_1 are close to the true value of β_1 , then the a priori auxiliary information should always contribute to reducing the mean square error (MSE) of β_1 .

If, however, the sample size is small, there is no apriori information linking the parameters of β_1 with β_2 , and in addition the correlation between the injury tolerance and the side effect (i.e., preinjury) is low, then the incorporation of preinjury data could possibly contribute to a slight increase in the MSE of the estimate of β_1 . In order to assess the benefits of the inclusion of preinjury data for small sample sizes, a Monte Carlo study is being conducted. This study will be discussed in a future report.

V. REFERENCES

- [1] Dixon, W. J., ed., BMDP: Biomedical Computer Programs, University of California Press, Berkeley, CA, 1977.
- [2] Helwig, J. and Council, K., eds., SAS User's Guide, SAS Institute, Inc., Raleigh, NC, 1979.
- [3] Peterson, J. J. and Smith, D. E., "A General Statistical Approach for Using Auxiliary Information in the Development of an Impact Acceleration Injury Prediction Model," Technical Report No. 112-6, Desmatics, Inc., August 1980.

APPENDIX

Since

$$\text{Var}(\epsilon_{21}y_{11}) = E\{(\epsilon_{21}y_{11})^2\} - \{E(\epsilon_{21}y_{11})\}^2,$$

it is sufficient, in view of the appendix in [3], to derive an expression for $E\{(\epsilon_{21}y_{11})^2\}$.

Now,

$$\begin{aligned} E\{(\epsilon_{21}y_{11})^2\} &= E\{\epsilon_{21}^2 y_{11}^2\} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \epsilon_{21}^2 I_1(t_1) f_{\epsilon, T}(\epsilon_{21}, t_1) d\epsilon_{21} dt_1 \\ &= \int_{-\infty}^0 \int_{-\infty}^{\infty} \epsilon_{21}^2 f_{\epsilon}(\epsilon_{21} | t_1) f_T(t_1) d\epsilon_{21} dt_1 \\ &= \int_{-\infty}^0 E\{\epsilon_{21}^2 | t_1\} f_T(t_1) dt_1 \end{aligned}$$

However,

$$\begin{aligned} E\{\epsilon_{21}^2 | t_1\} &= \text{Var}(\epsilon_{21} | t_1) + \{E(\epsilon_{21} | t_1)\}^2 \\ &= (1 - \rho^2)\sigma_2^2 + \rho^2\sigma_2^2(t_1 - \mu_{T_1})^2/\sigma_T^2 \end{aligned}$$

since $\epsilon_{21} | t_1$ has a normal distribution with mean

$$\rho\sigma_2(t_1 - \mu_{T_1})/\sigma_T$$

and variance

$$(1 - \rho^2)\sigma_2^2.$$

Thus,

$$E\{(\varepsilon_{2i}y_{1i})^2\} = (1 - \rho^2)\sigma_2^2 \int_{-\infty}^0 f_T(t_1) dt_1 + (\rho^2\sigma_2^2/\sigma_T^2) \int_{-\infty}^0 (t_1 - \mu_{T_1})^2 f_T(t_1) dt_1 .$$

Now, from (3) and (4) in the appendix of [3],

$$\int_{-\infty}^0 f_T(t_1) dt_1 = \int_{-\infty}^{-\mu_{T_1}/\sigma_T} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

$$= \Phi(-\mu_{T_1}/\sigma_T)$$

$$= \Phi(\underline{x}'_1 \underline{\beta}_1) .$$

Also,

$$\int_{-\infty}^0 [(t_1 - \mu_{T_1})^2/\sigma_T^2] f_{T_1}(t_1) dt_1 = \int_{-\infty}^{-\mu_{T_1}/\sigma_T} u^2 \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du .$$

By performing integration by parts, the integral on the right hand side becomes

$$-u(2\pi)^{-1/2} e^{-1/2 u^2} \Big|_{-\infty}^{-\mu_{T_1}/\sigma_T} + \int_{-\infty}^{-\mu_{T_1}/\sigma_T} \frac{1}{\sqrt{2\pi}} e^{-1/2 u^2} du$$

This expression is equal to

$$-(\mu_{T_1}/\sigma_T)\phi(-\mu_{T_1}/\sigma_T) + \Phi(\mu_{T_1}/\sigma_T) = -(\underline{x}'_1 \underline{\beta}_1)\phi(\underline{x}'_1 \underline{\beta}_1) + \Phi(\underline{x}'_1 \underline{\beta}_1)$$

Therefore,

$$E\{(\varepsilon_{2i}y_{1i})^2\} = (1 - \rho^2)\sigma_2^2 \Phi(\underline{x}'_1 \underline{\beta}_1) + \rho^2\sigma_2^2 [\phi(\underline{x}'_1 \underline{\beta}_1) - (\underline{x}'_1 \underline{\beta}_1)\phi(\underline{x}'_1 \underline{\beta}_1)] ,$$

and hence,

$$\text{Var}(\varepsilon_{2i}y_{1i}) = \sigma_2^2 [\phi(\underline{x}'_1 \underline{\beta}_1) - \rho^2 ((\underline{x}'_1 \underline{\beta}_1)\phi(\underline{x}'_1 \underline{\beta}_1) + \phi^2(\underline{x}'_1 \underline{\beta}_1))] .$$

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 112-7	2. GOVT ACCESSION NO. AD-A095299	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) COMPUTATIONAL ASPECTS OF INCORPORATING AUXILIARY INFORMATION INTO AN IMPACT ACCELERATION INJURY PREDICTION MODEL		5. TYPE OF REPORT & PERIOD COVERED Technical Report
7. AUTHOR(s) John J. Peterson and Dennis E. Smith		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Desmatics, Inc. P. O. Box 618 State College, PA 16801		8. CONTRACT OR GRANT NUMBER(s) N00014-79-C-0128
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Arlington, VA 22217		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR 207-037
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE February 1981
		13. NUMBER OF PAGES 17
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Distribution of this report is 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 Prediction Auxiliary Information Statistical Computation		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Auxiliary information may sometimes be used in development of a mathematical model in order to improve the estimated values of unknown parameters. This report discusses computational procedures which allow the application of commonly used nonlinear estimation programs to incorporate various sources of auxiliary information into an impact acceleration injury prediction model.		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

