To the Graduate Council:

I am submitting herewith a thesis written by Bradley W. Dolan entitled "A Statistical Method for Imputation of Uranium-235 Lung Burden Using Urinalysis Data." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Nuclear Engineering.

Peter G. Groër

Dr. Peter G. Groër, Major Professor

We have read this thesis and recommend its acceptance:

David R Simpson

L F Miller

Accepted for the Council:

C W Minkel

Associate Vice Chancellor and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature Bradley W. Dolan

Date December 31, 1991

A STATISTICAL METHOD FOR IMPUTATION OF URANIUM-235 LUNG BURDEN USING URINALYSIS DATA

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Bradley W. Dolan

May 1992

ACKNOWLEDGEMENTS

I am indebted to Dr. Peter G. Groër for the encouragement and assistance he provided while serving as my major professor. I also gratefully acknowledge the assistance of Nuclear Engineering Department Head Dr. Tom W. Kerlin and committee members Dr. Laurence F. Miller and Dr. David R. Simpson.

In addition, I express my gratitude to Dr. Donna L. Cragle, Mr. Charles M. West, Mr. William G. Tankersley, and the rest of the staff of the Center for Epidemiologic Research for their invaluable support.

Data for this research were drawn from the Oak Ridge Associated Universities' Center for Epidemiologic Research Data Model, which is maintained to support the ORAU Health and Mortality Studies of Department of Energy Workers. This work was done under contract No. DE-AC05-76OR00033 between ORAU and the Department of Energy, by agreement between ORAU and the University of Tennessee.

ABSTRACT

Epidemiologic studies of groups of radiation workers are being performed by the Oak Ridge Associated Universities' Center for Epidemiologic Research in order to identify radiation-induced health effects and to quantify potential dose-response relationships. The historical radiological monitoring data sets available for a number of groups under study do not directly provide sufficient information to permit calculation of workers' doses. An example of this problem has been encountered with the Fernald Feed Materials Production Center studies. Biological monitoring data for workers exposed to airborne uranium dust in operations there consists of a large set of urinalysis records reporting mass quantities of uranium and a small set of lung count records reporting quantities of Uranium-235 and Uranium-238 present in the lung, estimated by gamma spectroscopy.

Bayesian linear regression and prediction techniques have been applied to this data to enable the prediction of lung burdens of Uranium-235 in workers based on their urinalysis results and a "calibration" data set of urinalysis results and measurements of Uranium-235 lung activity. The output of the procedure is a Uranium-235 lung burden estimate provided as a probability density. Computer

programs have been written to facilitate large-scale application of the procedure.

As the precision of the input urinalysis data was not known, an exploration was made of the effects of uncertainty in this data set on regression parameters and prediction . A computer program was written to simulate the effect of uncertainty in the urinalysis measurements. A random value was chosen from an assumed distribution about each urinalysis result in the calibration data set and a representative urinalysis result for which a prediction was desired. The regression and prediction procedures were then performed, and this process was repeated a number of times. Distributions of results were obtained which indicate the effects of uncertain urinalysis values on regression parameters and prediction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

1.1 Epidemiologic Studies

Fundamental questions underlying the field of radiation protection are: "What effects do ionizing radiations have on humans?", and "What are the quantitative relationships between radiation doses and effects?" The process of radioactive decay is considered stochastic, as are the biological late effects of ionizing radiations (Turner 1986, BEIR V 1990). To date, no deterministic method has been identified to predict the human health effects which will occur as a result of exposures to low doses of ionizing radiations and it seems unlikely that a means for deterministic prediction of these effects will be found.

An approach to answering the above questions has been to perform bio-statistical studies of groups of persons exposed to radiation, to identify and quantify relationships between radiation exposures, and health outcomes. Several such studies are currently in progress at the Oak Ridge Associated Universities' Center for Epidemiologic Research (CER).

One study in progress there will investigate the lung cancer mortality experienced by workers occupationally exposed to uranium dust. This study attempts to identify relationships between mortality from lung cancer and radiation dose to the lung. The ideal dosimetry information for such a study would be the lung dose from uranium for each worker, as a function of time. Biological monitoring data actually available, however, do not permit calculation of this lung dose. The data available consist primarily of a large body of urinalysis records in mass units of uranium without regard to activity, and a small body of lung counts reporting estimated quantities of Uranium-235 and Uranium-238 in the lung.

1.2 Radiological Monitoring Data Available

A series of uranium lung activity measurements, taken at regular intervals over the whole employment period, would provide sufficient information about workers' exposures to radiation from uranium deposits in the lung. Weighted by time, these activity measurements could be used directly in an epidemiologic analysis or could provide the basis for the estimation of lung doses for use in the study.

This level of dosimetric information desired for epidemiologic studies is rarely encountered in practice.

Illustrative of the data typically encountered is the biological monitoring data set from the facility under study, the Fernald Feed Materials Production Center (FMPC), a uranium processing and refining plant.

This facility was constructed in 1951 by an Atomic Energy Commission contractor to process uranium ore concentrates and other uranium materials of relatively low enrichments into fabricated metal products such as reactor core target elements for the Savannah River Plant (Fugate 1982).

From the inception of operation through 1983, the last year for which complete exposure data was available at CER, workers were monitored for external beta and gamma radiation doses using film badges and the annual summary results were reported. A urinalysis program which reported mass units of uranium (micrograms/liter) provided the primary indication of workers' internal exposures from 1952 through 1967. A very small body of urinalysis records reporting activity units is believed to exist but was not available in the CER computerized database. As table I-1 indicates, the average urinalysis frequency for monitored workers varied from 7.4 / year in 1954 to 2.4 / year in 1969.

Beginning in 1968, the urinalysis program was augmented by a lung-counting program. This program applied gamma spectroscopy techniques to estimate quantities of

Table I-1. Urinalysis Frequency

Year	Total Urinalyses	Monitored Workers*	Average Sampling Frequency
1952	257	56	4.6 per year
1953	3454	770	4.5
1954	9395	1278	7.4
1955	10935	1891	5.8
1956	10551	2343	4.5
1957	13506	2851	4.7
1958	12807	2509	5.1
1959	14409	2602	5.5
1960	19237	2695	7.1
1961	9450	2431	3.9
1962	8424	2175	3.9
1963	9608	2022	4.8
1964	6707	1905	3.5
1965	6182	1670	3.7
1966	5722	1608	3.6
1967	5581	1666	3.3
1968	4599	1414	3.3
1969	3152	1291	2.4
1970	3005	1109	2.7
1971	2229	879	2.5
1972	2175	658	3.3
1973	2667	728	3.7
1974	2452	676	3.6
1975	2261	684	3.3
1976	1996	667	3.0
1977	1805	638	2.8
1978	1621	615	2.6
1979	1537	583	2.6
1980	1556	606	2.6
1981	2027	651	3.1
1982	2838	818	3.5
1983	3277	870	3.8

*Monitored worker defined as worker on urinalysis program.

Source: Database, Center for Epidemiologic Research, Oak Ridge Associated Universities.

Uranium-235 and Uranium-238 present in workers' lungs. As table I-2 indicates, the fraction of monitored workers given lung counts varied from 0.76 / year in 1971, to 0.08 / year in 1969.

Data of a type well suited for use in an epidemiologic study, estimated lung burdens of Uranium-235 and Uranium-238, exist for some workers after 1967. For the first sixteen years of plant operations, however, only urinalysis data exist to indicate internal exposures. In addition, little or no lung count data is available for many other exposed workers employed since the inception of the lung-counting program.

Epidemiological analysis of the worker population would be greatly facilitated if a method could be developed to predict lung burdens of uranium using existing data. As one or more urinalysis measurements exist in nearly all years of employment for nearly all production workers, an approach utilizing the urinalysis data to infer uranium lung activity is suggested. Use of existing methods such as metabolic models to estimate lung burden or related quantities using the urinalysis data is problematic, though, without information regarding such variables as exposure patterns, material enrichment and solubility.

Table I-2. Lung Count Frequency

Year	Total Lung Counts	Monitored Workers	Fraction Counted
1968	314	1414	0.22 per year
1969	97	1291	0.08
1970	153	1109	0.14
1971	669	879	0.76
1972	259	658	0.39
1973	209	728	0.29
1974	302	676	0.45
1975	249	684	0.36
1976	228	667	0.34
1977	206	638	0.32
1978	197	615	0.32
1979	184	583	0.32
1980	199	606	0.33
1981	146	651	0.22
1982	180	818	0.22
1983	192	870	0.22

Source: Database, Center for Epidemiologic Research, Oak Ridge Associated Universities.

1.3 Review of Related Research

Other researchers have addressed related problems involving inference of certain bioassay measurements for uranium workers using other bioassay measurements or air sampling measurements.

Boback (1975) described instances of elevated urinalysis measurements for workers at the FMPC. He compared urinalysis results with air sampling data in one case and noted that "no significant exposure conditions were disclosed by air sampling."

Scott and West (1975) examined urinalysis and lung count data for uranium workers at the Y-12 weapons plant and found "no definitive relationship between the two programs," observing that "the lack of correlation can be partially explained by the fact that neither program is very precise." Somewhat in contradiction, they also reported that, "if the urine average is high, the in vivo average tends to be high also" for employees working in the same area.

More recently, Spitz, et. al. (1984) compared urinalysis and in vivo measurements for workers at several uranium mills. Their finding was that "urinalysis results were highly variable and thus unreliable as a monitoring tool."

Smith et. al. (1987) compared urinalysis and in vivo measurements for workers involved in an exposure incident at the FMPC. They reported that urinalysis results for about one hundred workers did not show evidence of internal uranium contamination but that lung counts for two of the workers indicated uranium body burdens slightly above the minimum detectable level.

Chase has compared air monitoring and urinalysis data from workers at a uranium mill (1989), finding that "U(ranium) urinalysis results correlate with airborne U exposure levels when averaged over a period of time." He found that average airborne uranium activity levels were related to average group urinalyses, and that trends in airborne activity levels resulted in similar trends in individual workers' urinalysis results. He did not attempt to relate airborne activity or urinalysis measurements to lung activity measurements.

1.4 Contribution of Thesis

Prior efforts to estimate lung counts based on other monitoring information have done little with respect to quantifying relationships between lung counts and air monitoring or urinalysis data, or other measures.

This thesis demonstrates the application of Bayesian linear regression and prediction techniques to impute the lung burden of Uranium-235 present in workers in the FMPC study population, using the more generally available urinalysis measurements in conjunction with a data set of paired urinalysis and lung counts. The technique presents estimates as probability density functions, which convey the uncertainty associated with the estimation procedure. Computer programs to perform the regression and prediction procedures are described. The output of the predictive program may be used as input to epidemiologic studies of the Fernald worker population. These programs can also be used for other studies with similar exposure data problems.

In addition, an exploration has been made of the effects of uncertain urinalysis data on the prediction of Uranium-235 lung activity. A computer program simulating the uncertainty of the input urinalysis records is provided and the results of the simulation study are discussed. The results of this investigation are of use in determining the precision of radiation risk estimates provided by epidemiologic studies.

1.5 Organization of Thesis

Following the introduction are these chapters: "Fernald Feed Materials Production Center Exposure Hazards and Radiological Monitoring Data", "Imputation of Uranium-235 Lung Burden", "Effect of Uncertain Urinalysis Data", and "Discussion and Suggestions for Further Research." Appendixes providing the calibration data set employed in example calculations, details concerning the derivation of the regression parameter and predictive distributions, listings of the computer programs utilized, and sample output from the programs follow.

In the chapter on the radiological monitoring data, the processes utilized at the FMPC are described to provide understanding of the nature of the exposure hazards experienced by workers. Details of the radiological monitoring practices are given and the resulting data sets are described.

In the chapter on the imputation procedure, the Bayesian linear regression technique is described and a description of the derivation of the posterior distributions of the regression parameters is given. The derivation of the predictive distribution follows. The computer programs to execute different statistical procedures are discussed. Sample input and output from these programs are presented.

The chapter describing the investigation of the influence of uncertain monitoring data tells how the uncertainty of the urinalysis data in the FMPC data set was simulated with a computer program and provides insights gained regarding the accuracy of regression and prediction parameters estimated using uncertain data.

Finally, the findings of the thesis are summarized in the final chapter and suggestions for future research are made.

CHAPTER II

FERNALD FEED MATERIALS PRODUCTION CENTER EXPOSURE HAZARDS AND RADIOLOGICAL MONITORING DATA

2.1 Processes and Exposure Hazards at the Fernald Feed Materials Production Center

In order to understand the radiation protection concerns relevant to the workers at the Fernald FMPC, it is helpful to examine the production processes employed there. The FMPC was constructed near Fernald, Ohio in 1951. It was operated by the National Lead of Ohio Company as a contractor to the Atomic Energy Commission and its successors. In the mid-nineteen-eighties, management of the facility was transferred to the Westinghouse Materials Company of Ohio. Designed to refine, smelt, and machine low-enrichment uranium products, the Fernald FMPC employed over 5,000 workers between 1951 and 1983, the last year for which complete data is available at the CER. Today, work at the Fernald FMPC is comprised of cleanup and environmental restoration efforts.

The diverse production operations there were spread among 9 main facilities, or plants (Fugate 1982). Plant 1, the sampling plant, received, sampled, and stored large quantities of uranium materials intended for subsequent processing. Workers in the plant measured uranium content and enrichment in incoming materials, typically arriving in drums, and crushed and milled them as required for use in downstream processing. A small safe-geometry digestion system was installed in the plant in 1970 to permit processing of limited quantities of uranium enriched up to 10% in Uranium-235.

Plants 2 and 3 comprised the refinery operation. There, ore concentrates and recyclable materials were digested in nitric acid and subjected to solvent extraction steps involving tributyl phosphate, kerosene, and water. The purified uranium solution was calcined in denitration pots to yield UO_3. Equipment was also installed in 1972 to convert up to 10% enriched Uranium-235 solution to U_3O_8, "black oxide".

Plant 4 converted UO_3 to UF_4, "green salt", in a continuous flow chemical reactor system. UO_3, also referred to as orange oxide, was reduced with hydrogen to form UO_2 and then reacted with hydrogen fluoride to to yield UF_4. The UF_4 was packaged in 10-gallon cans for shipment to other processing sites or for use in processes at Fernald.

Workers in the metals production plant, plant 5, tamped UF_4 and magnesium in reduction pots which were then heated to reduce the mixture to uranium metal "derbies" and slag. Slag was removed from the derbies by use of chipping hammers and shot-blasting with uranium shot metal. Cleaned derbies were then remelted to form ingots which were cropped and sawed to desired dimensions.

Plant 6, the metals fabrication plant, rolled the ingots into rods and also cut and heat-treated fabricated metal pieces received from the onsite special products plant and offsite extrusion facilities. Machining chips were crushed, cleaned, and briquetted for recycling.

Scrap recovery was performed in plant 8, where recycle materials were cleaned, classified, treated as required, and returned to appropriate points in the plant process stream. Only low-enrichment uranium materials, less than 1.25% Uranium-235, could be processed there.

The special products plant, plant 9, machined ingots and billets for extrusion. Ingots were cropped, center-drilled and surface machined with lathes. Chemical decladding of imperfect new reactor fuel assemblies was also performed in plant 9. Equipment was installed in the special products plant to permit reduction of ThF_4 to thorium and to machine the thorium, however little information is available about its use.

Information was also limited concerning plant 7, which was a thorium processing facility. This plant was apparently only operated for a short time, a number of years ago. The pilot plant on site also was equipped to purify and handle thorium materials. A safe-geometry purification facility for uranium enriched up to 10% was installed in the pilot plant in 1976. Recycle materials enriched between 1.25% and 10% Uranium-235 were cleaned at the pilot plant and returned to the process stream. Numerous other operations were conducted in the pilot plant, including processing of uranium hexaflouride (Boback and Heatherton 1966).

Thus, it is evident that workers at the Fernald plant were exposed to uranium in numerous chemical and physical forms. Little information is available regarding the specific forms to which each individual worker may have been exposed, the sizes of the airborne particles involved, etc. Because of this, as well as limitations in the quantity of data available, application of metabolic models to estimate workers' lung doses would be difficult.

2.2 Radiological Monitoring Practices

Production workers at Fernald were monitored for external radiation exposure. The external dosimetry program measured worker's beta and gamma radiation doses on a monthly basis using film badges. External dose totals in rad were reported annually.

The primary biological monitoring technique employed at the FMPC was urinalysis for uranium. Employing a fluorometric technique (Boback 1966; Smith 1987), the program reported mass units (micrograms) per liter of urine. Specifics of the Fernald program are not readily available, but the literature provides general information about fluorometric urinalysis (Baselt 1988). Typically, an aliquot of urine is dried and fused with sodium fluoride. This mixture is subjected to an ultraviolet light and the resulting fluorescence is measured. The light emission of a urine blank is subtracted and the difference is compared to the light emissions of known standard samples processed in the same way. Baselt indicates that the minimum level of uranium detectable by this procedure is about 5 micrograms per liter. He also indicates that measurement variability will be reduced by sampling workers on return to work after a weekend and it is believed that the FMPC followed this practice. This practice also allows high-solubility uranium compounds which contribute little to dose to be eliminated before sampling. Table II-1 depicts average urinalysis values by year for FMPC workers.

Table II-1. Annual Mean Urinalysis and Lung Count Values

Year	Mean Urinalysis (Micrograms per liter)	Mean Lung Count U-235 (micrograms)	Mean Lung Count U-238 (milligrams)
1952	80	–	–
1953	74	–	–
1954	44	–	–
1955	56	–	–
1956	39	–	–
1957	181 *	–	–
1958	31	–	–
1959	18	–	–
1960	23	–	–
1961	20	–	–
1962	15	–	–
1963	15	–	–
1964	16	–	–
1965	12	–	–
1966	19	–	–
1967	8	–	–
1968	8	74	12
1969	7	70	20
1970	5	73	13
1971	7	44	8
1972	16	39	11
1973	9	39	5
1974	7	75	16
1975	8	52	8
1976	7	47	6
1977	6	30	6
1978	6	45	9
1979	7	50	13
1980	6	48	8
1981	4	9	14
1982	4	23	5
1983	5	14	-1.3

*Mean urinalysis in 1957 was strongly influenced by extreme values. Mean of records <2000 micrograms/L was 43.

Source: Database, Center for Epidemiologic Research, Oak Ridge Associated Universities.

Beginning in 1968, the Y12 mobile in-vivo radiation monitoring laboratory ("lung counter") was used to estimate uranium lung burdens of some workers. This lung counter utilized two thallium-activated sodium iodide crystals, 9 inches in diameter and 4 inches thick. Four photomultiplier tubes converted the light emissions of each crystal to electrical signals. A five-inch thick iron shield was employed to reduce background radiation counts (Anonymous 1968).

Workers were counted for twenty minutes in the lung counter, with counts about the 90, 186, 662, and 1430 keV regions summed separately (Scott and West 1975; Smith et. al. 1987). The 186 keV region observed the Uranium-235 gamma emission and the 90 keV region observed the Th-234 90 keV X-ray. Thorium-234 is a daughter product of Uranium-238 and, if the two are in equilibrium, one can be used to infer the quantity of the other. Counts in the 662 keV and 1430 keV regions were evaluated to assess respectively the Compton scatter contribution from Cesium-137 and Potassium-40 to the counts in the Uranium-235 and Thorium-234 regions. After subtraction of background contribution, these counts were used in a prediction equation to estimate the quantities of Uranium-235 and/or Uranium-238 present. (Personal Communication, 1988) Occasionally, counts were reported for other materials.

Smith et. al. (1987) estimated that the minimum detectable level of Uranium-235 in this counter was 49 micrograms and the minimum detectable level of Uranium-238 was 5.0 milligrams. They concluded this based on observing counts for a number of unexposed persons. The distribution of estimated Uranium-235 lung burdens for this group was centered at 0 and had a standard deviation of 24.5 micrograms. The distribution of estimated Uranium-238 burdens was also centered at 0 with a standard deviation of 2.5 milligrams. Assuming that the distributions are approximately normal, estimated lung burdens more than two standard deviations above the mean would only be observed about 2 percent of the time in unexposed persons. Therefore, estimated lung burdens greater than 49 micrograms Uranium-235 or 5.0 milligrams Uranium-238 were considered to indicate the presence of uranium.

The mobile counter visited the FMPC site once or twice a year and would typically count a few hundred workers. These were primarily those believed to have higher potential for exposure (Heatherton 1975). Ordinarily, about a third of the workers monitored on the urinalysis program would receive lung counts in a given year. Average estimated Uranium-235 and Uranium-238 lung burdens are shown, by year, in Table II-1.

Airborne uranium activity was also measured in the various production areas. One set of measures were apparently taken in potential problem areas, another set of measures were taken to try to represent average worker exposures in particular plants. Neither set of records provides any information about the specific workers who were exposed to the air concentrations measured and therefore neither can be used for dosimetry.

2.3 Personnel Monitoring Data Sets

Personnel radiation monitoring data sets from the FMPC consist of annual film badge external dose totals, records of urinalysis for mass uranium, and lung counts. Each of these data sets is maintained as a computer file at the CER.

The external dosimetry program provided the most comprehensive coverage of the workforce, but does not appear relevant in assessing the radiological hazard of greatest interest, exposure to airborne uranium dust.

The lung counting data consists of about 5,000 records, a few hundred for each year since 1968. Most records report estimated lung burden of Uranium-235 and Uranium-238. The number of records occurring in each year is identified in Table I-2 on page 5, and the mean quantities of Uranium-235 and Uranium-238 measured are identified in Table II-1.

Approximately 185,000 urinalysis records provide information about mass quantities of uranium present in workers' urine. A copy of this data set is maintained at the CER and spans the periods from 1952 to 1983. Annual record counts for this data are given in Table I-1 on page 5 and average urinalysis values are summarized in Table II-1.

CHAPTER III

IMPUTATION OF URANIUM-235 LUNG BURDEN

3.1 Construction of Calibration Data set

With regard to epidemiologic studies, the difficulty with the Fernald radiation monitoring data set is that the most useful exposure measure, lung counts, are available only for some workers and only for some time periods. A less useful measure, urinalysis, exists for all years of plant operation and for most exposed workers. The problem, then, is to use a relationship between the two measures so that one may be used to predict the other.

After 1967, estimates of uranium lung burden were made for some workers at the FMPC. Typically, lung burdens were reported as estimated milligrams of Uranium-238 and micrograms of Uranium-235 present. The quantity of Uranium-235 present was estimated directly by counting the Uranium-235 180 keV gamma emission. The quantity of Uranium-238 present was inferred by counting the 90 keV Thorium-234 X-ray and assuming that the daughter nuclide was in equilibrium with the parent. Scott and West (1975) have indicated that this method is unreliable for facilities which melt and process molten uranium, reporting the following: "The daughter melts and floats to the surface, thereby enriching the surface of the melted part. False high results sometimes are obtained on employees from this type operation." They also note that "if an exposure occurs to freshly purified uranium, it could go undetected due to the absence of the daughter." As much of the work at the FMPC entailed melting of uranium or processing of recently melted uranium products, the reliability of the Uranium-238 lung burden estimates appears suspect. Therefore, in this thesis, we examine only the data pertaining to Uranium-235 lung burden. Given appropriate assumptions about enrichment, this measure may be used to infer the quantity of Uranium-238 present. The quantities of Uranium-234 and uranium daughter nuclides present may also be inferred.

Simultaneous urinalyses and Uranium-235 lung count results can be used as a calibration data set, $D = \{X_i, Y_i, i = 1, 2, ..., n\}$, for imputations of lung burdens in workers for whom only urinalysis results are available. If Y_N denotes the lung burden of Uranium-235 and X_N, a new urinalysis result for an individual for whom Y_N is not measured, then a procedure is needed to infer a probability distribution for Y_N based on the observation of X_N and the calibration data, D. Symbolically, the predictive density of $p(Y_N | X_N, D)$ is desired.

A calibration data set was constructed using the FMPC urinalysis and lung counting computer data files maintained in the Center for Epidemiologic Research database. Approximately five thousand lung count records exist for Fernald workers. Of these lung counts, approximately two thousand were found to occur in the same calendar month as a urinalysis record for the same worker. These matching records were paired and written to a computer file. A subset of these paired records, those with urinalyses and lung counts occurring within one week of each other, was examined to see if the two variables would be more closely related. Data for each plant were examined separately. A large reduction in the quantity of data resulted, but inspection of scatterplots and comparison of correlation coefficients did not identify a closer relationship between the two variables. The larger data set of records occurring in the same month was therefore retained.

As previously described, different chemical and metalurgical processes were performed in each plant at the FMPC. Each plants' processes gave rise to differing exposure hazards; based on the level of airborne activity, particle sizes, enrichments, material solubility, etc. It seems reasonable that a plant-specific calibration data set would enable better predictive precision for workers in that plant than could be obtained by using data for all FMPC workers. For this reason, it was decided to focus on data from workers in one plant, plant 8. The exposure hazards in the scrap recovery plant appear to be representative of a number of the hazards encountered throughout the FMPC. There were numerous chemical and metallic forms of uranium handled there; processes employed included crushing, milling and furnacing. The enrichment of materials handled there is somewhat better known, than for other plants, however, as only materials of less than 1.25% enrichment were handled in the recovery plant.

Restricting the calibration data set to records for plant 8 workers yielded a data set of 102 records. Eleven of the records reported negative quantities of Uranium-235 present. Since this is not possible, these records were adjusted to show 0 milligrams of Uranium-235. This data set, which was employed for the analyses subsequently described, is listed in Appendix A. Further discussion concerning the handling of reported negative dosimetric data appears in Appendix D.

3.2 Posterior Distributions of the Regression Parameters

If it is desired to find the distribution of a parameter (or vector of parameters), $\boldsymbol{\theta}$, and data is available whose distribution is dependent on $\boldsymbol{\theta}$, then one approach to estimating θ is through the application of Bayes' theorem. If "DATA" denotes observations, $p(DATA|\theta)$ denotes the probability of the data conditional on θ, and the prior density $p(\theta)$ denotes the distribution of θ prior to incorporation of the data; then a posterior density function $p(\theta|DATA)$ may be determined which describes the uncertainty about θ after the data have been incorporated:

$$p(\theta|DATA) = \frac{p(DATA|\theta)\ p(\theta)}{\int_{-\infty}^{\infty} p(DATA|\theta)\ p(\theta)\ d\theta}$$

Since the denominator is not dependent on θ, this expression is generally rewritten:

$$p(\theta|DATA) \propto p(DATA|\theta)\ p(\theta)$$

The above relationship, Bayes' theorem, is applied in this thesis to estimate the parameters of a linear regression of lung count measurements on urinalyses and to derive an expression for a predictive distribution. The linear regression model is given by:

$$Y_i|X_i = B_0 + B_1X_i + \mu_i, \quad i=1,2,\ldots,n$$

where X_i and Y_i represent the i-th observation on independent and dependent variables respectively. The random errors of Y_i are represented by μ_i. For simplicity, these errors are assumed to be independent and identically distributed as $N(0,\sigma^2)$. The unknown parameters of the model, then, are B_0, B_1, and σ. These parameters must be estimated from the data.

The likelihood function in this case is given:

$$p(\underline{Y},\ \underline{X} \mid B_0,\ B_1,\ \sigma) \propto \frac{1}{\sigma^n}\, e^{\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (Y_i-B_0-B_1X_i)^2\right\}}$$

When little is known about the possible values of the model parameters prior to examining the data, Press (1989) suggests the adoption of a non-informative prior density:

$$p(B_0) \propto \text{CONSTANT} \qquad p(B_1) \propto \text{CONSTANT} \qquad p(\sigma) \propto 1/\sigma$$

$$0 < \sigma < \infty, \qquad 0 \leq B_0 < \infty, \qquad 0 < B_1 < \infty$$

$$p(B_0,\ B_1,\ \sigma) = p(B_0)p(B_1)p(\sigma) \propto 1/\sigma$$

This implies that the probabilities of $\widehat{B_0}$ and $\widehat{B_1}$ falling into any interval of fixed size between 0 and $+\infty$ are assumed equal, while large values of σ are less likely than smaller ones. The joint posterior density is obtained by multiplying the likelihood and the prior:

$$p(B_0, B_1, \sigma \mid \underline{X}, \underline{Y}) \propto \frac{1}{\sigma^{n+1}} e^{\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (Y_i - B_0 - B_1 X_i)^2\right\}}$$

The maximum likelihood estimators, $\widehat{B_0}$ and $\widehat{B_1}$, are computed by minimizing the sum of the squares of the residuals, where the residuals are the errors about the regression line (Details of this and other results referenced in this section are provided in Appendix B):

$$\widehat{B_0} = \frac{\sum_{i=1}^{n} Y_i - B_1 \sum_{i=1}^{n} X_i}{n}$$

$$\widehat{B_1} = \frac{n \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i\right)^2}$$

The posterior density of the parameters B_0 and B_1 may be obtained by integrating σ out of the joint posterior density and the result may be again integrated to obtain the marginal densities of B_0 and B_1, which are both t-distributions with n-2 degrees of freedom:

$$\left[\frac{\sum_{i=1}^{n} (X_i-\bar{X})^2}{\frac{s^2}{n}\sum_{i=1}^{n} X_i^2}\right]^{1/2} \times (B_0 - \widehat{B_0}) \mid \underset{\sim}{Y}, \underset{\sim}{X} \sim t_{n-2}$$

$$\frac{\left(\sum_{i=1}^{n} (X_i-\bar{X})^2\right)^{1/2}}{s} \times (B_1 - \widehat{B_1}) \mid \underset{\sim}{Y}, \underset{\sim}{X} \sim t_{n-2}$$

An unbiased estimator for σ^2 is (Walpole and Myers, 1989):

$$s^2 = \frac{\sum_{i=1}^{n} \left(Y_i - \widehat{B_o} - \widehat{B_1}X_i\right)^2}{n - 2}$$

The posterior densities of B_0 and B_1 were calculated for the sample data set by the FORTRAN program "CALC" (Appendix C1) written for this purpose.

$$\text{MODE}(B_0 \mid \underset{\sim}{Y}, \underset{\sim}{X}) = \widehat{B_0} = 45.8$$

$$\text{SD}\left(B_0 \mid \underset{\sim}{Y}, \underset{\sim}{X}\right) = 7.20$$

$$\text{MODE}(B_1 \mid \underset{\sim}{Y}, \underset{\sim}{X}) = \widehat{B_1} = 1.22$$

$$\text{SD}\left(B_1 \mid \underset{\sim}{Y}, \underset{\sim}{X}\right) = 0.51$$

The posterior densities may then be represented:

$$\frac{\left(B_0 - 45.80\right)}{7.20} \mid \underset{\sim}{Y}, \underset{\sim}{X} \sim t_{100}$$

$$\frac{\left(B_1 - 1.22\right)}{0.51} \mid \underset{\sim}{Y}, \underset{\sim}{X} \sim t_{100}$$

These distributions are displayed in figures III-1 and III-2. They are essentially normal in shape, resulting from the large number of data points (N=102). By inspection of the figures, it can be observed how well the regression parameters are known.

The posterior density distributions of the regression parameters may be used to make probability statements about the parameters. The 90% probability interval for B_0, for instance, is given:

$$p[(45.80 - 1.64 * 7.20) \leq B_0 \leq (45.80 + 1.64 * 7.20)] = .9$$

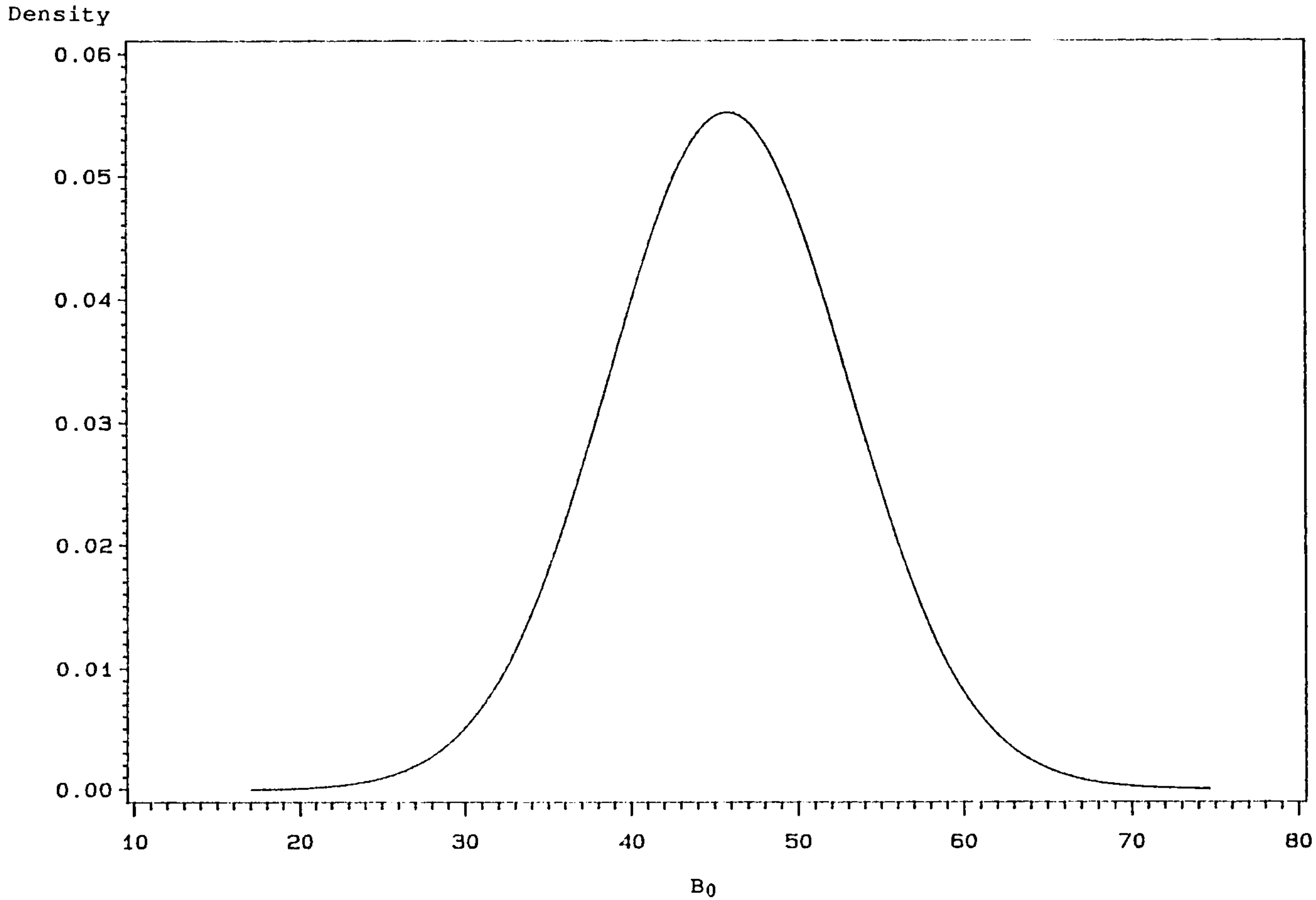

FIGURE III-1 Posterior Density of B_0. (Mean 45.8, SD 7.2, DF 100)

Micrograms U-235

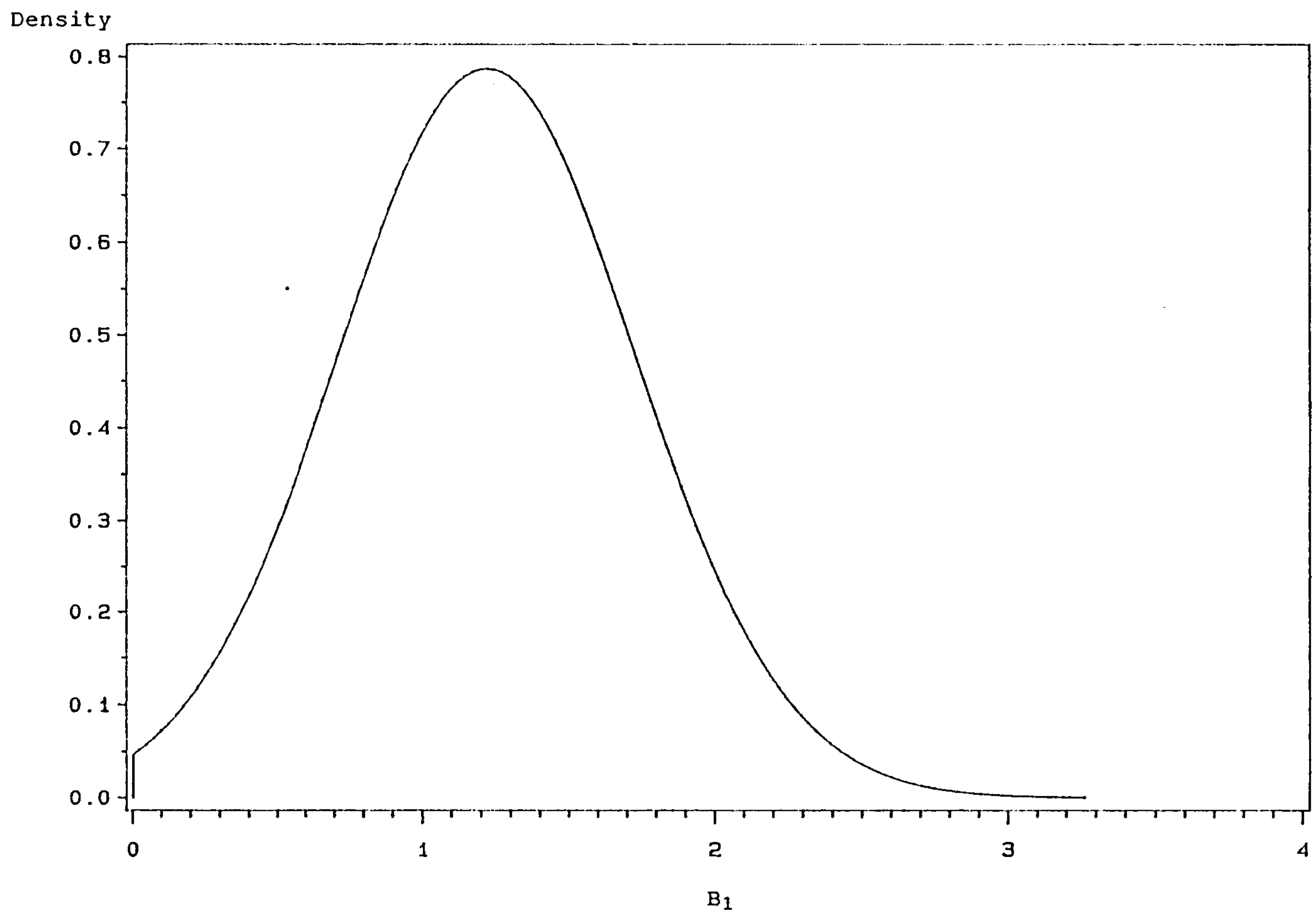

FIGURE III-2 Posterior Density of B_1. (Mean 1.22, SD 0.51, DF 100) Micrograms U-235 per Microgram/Liter

$$P(34.0 \leq B_0 \leq 57.6) = .9$$

By integrating B_0 and B_1 out of the joint density one obtains the posterior density for σ:

$$p\ (\sigma \mid \underline{X},\ \underline{Y}) = \int_0^\infty \int_0^\infty p(B_0,\ B_1,\ \sigma \mid \underline{X},\ \underline{Y})\ dB_0\ dB_1$$

$$\propto \frac{1}{\sigma^{n-1}}\ e^{-(n-2)\ s^2\ /2\sigma^2} \qquad \left(0 < \sigma < \infty\ ;\ n > 2\right)$$

Recognizing this as an inverted gamma distribution, the expectation and variance of σ can be written (Press 1989):

$$E\ (\sigma \mid \underline{X},\ \underline{Y}) = \frac{\Gamma\left(\frac{n-3}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} \left(\frac{n-2}{2}\right)^{1/2} s$$

$$SD\ (\sigma \mid \underline{X},\ \underline{Y}) = \left(\frac{(n-2)s^2}{(n-4)} - \left[E(\sigma \mid \underline{X},\ \underline{Y})\right]^2\right)^{\frac{1}{2}}$$

Using the "CALC" program, these values were calculated using the sample data set:

$$E\ (\sigma \mid \underline{X},\ \underline{Y}) = 46.3 \qquad SD\ (\sigma \mid \underline{X},\ \underline{Y}) = 3.3$$

3.3 Predictive Distribution for Uranium-235 Lung Burden

Using the calculated parameters and the calibration data set, it is possible to determine predictive intervals for lung counts imputed from new urinalysis data. If a new urinalysis measurement, X_N, is provided, the relation between Y_N and X_N given by the model is:

$$Y_N \mid X_N = B_0 + B_1 X_N + \mu \qquad \mu = N(0, \sigma^2)$$

The predictive density of Y_N is:

$$p\left(Y_N \mid X_N, \underline{Y}, \underline{X}\right) =$$

$$\int_0^\infty \int_0^\infty \int_0^\infty p\left(Y_N \mid X_N, B_0, B_1, \sigma\right) p\left(B_0, B_1, \sigma \mid \underline{X}, \underline{Y}\right) dB_0 \, dB_1 \, d\sigma$$

This integral is evaluated to yield the following expression for the density of Y_N:

$$\frac{Y_N - \widehat{B_0} - \widehat{B_1} X_N}{s\left[1 + \frac{1}{n} + \frac{\left(X_N - \overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\right]^{\frac{1}{2}}} \mid \underline{X}, \underline{Y} \sim t_{n-2}$$

The FORTRAN program "PREDICT" (Appendix C2) was written to perform the necessary calculations. Using the calibration data set and a test record with known lung count, Y_N',

$$X_N = 6 \text{ micrograms/liter} \quad , \quad Y_N' = 33 \text{ micrograms},$$

the following predictive distribution resulted:

$$\frac{Y_N - 53.1}{46.3} \sim t_{100}$$

The 90% probability interval is given:

$$p\left[(53.1 - 46.3 * 1.15) \leq Y_N \leq (53.1 + 46.3 * 1.37)\right] = .9$$
$$= p(0 \leq Y_N \leq 117.0)$$

The actual Uranium-235 lung burden value of $Y_N' = 33$ micrograms, which was known for this test case, falls well within this interval. This example predictive distribution is graphically displayed in figure III-3.

Figure III-4 illustrates a further application of the predictive distribution. The "PREDICT" program was executed using as input the plant-wide mean urinalysis value in 1952, 80 micrograms/L. Figure III-4 then illustrates the predicted Uranium-235 lung burden for a hypothetical, "average", worker whose urinalysis result was equal to the 1952 average. The distribution is centered at 143 micrograms and has a standard deviation of 58.2.

3.4 Computer Programs

Two computer programs were written in FORTRAN to facilitate imputation of uranium lung burdens by the methods described. The first, "CALC", reads a user-furnished calibration data file, calculates the necessary parameter estimates, and writes them to an output file. The second program, "PREDICT", reads the parameter estimates from this file and then operates on a user-furnished urinalysis input data set to generate lung burden predictive density distributions. The modes and standard deviations of the predictive density distributions are then written to an output file for subsequent use. Sample output based on the sample data set are included in Appendix C4.

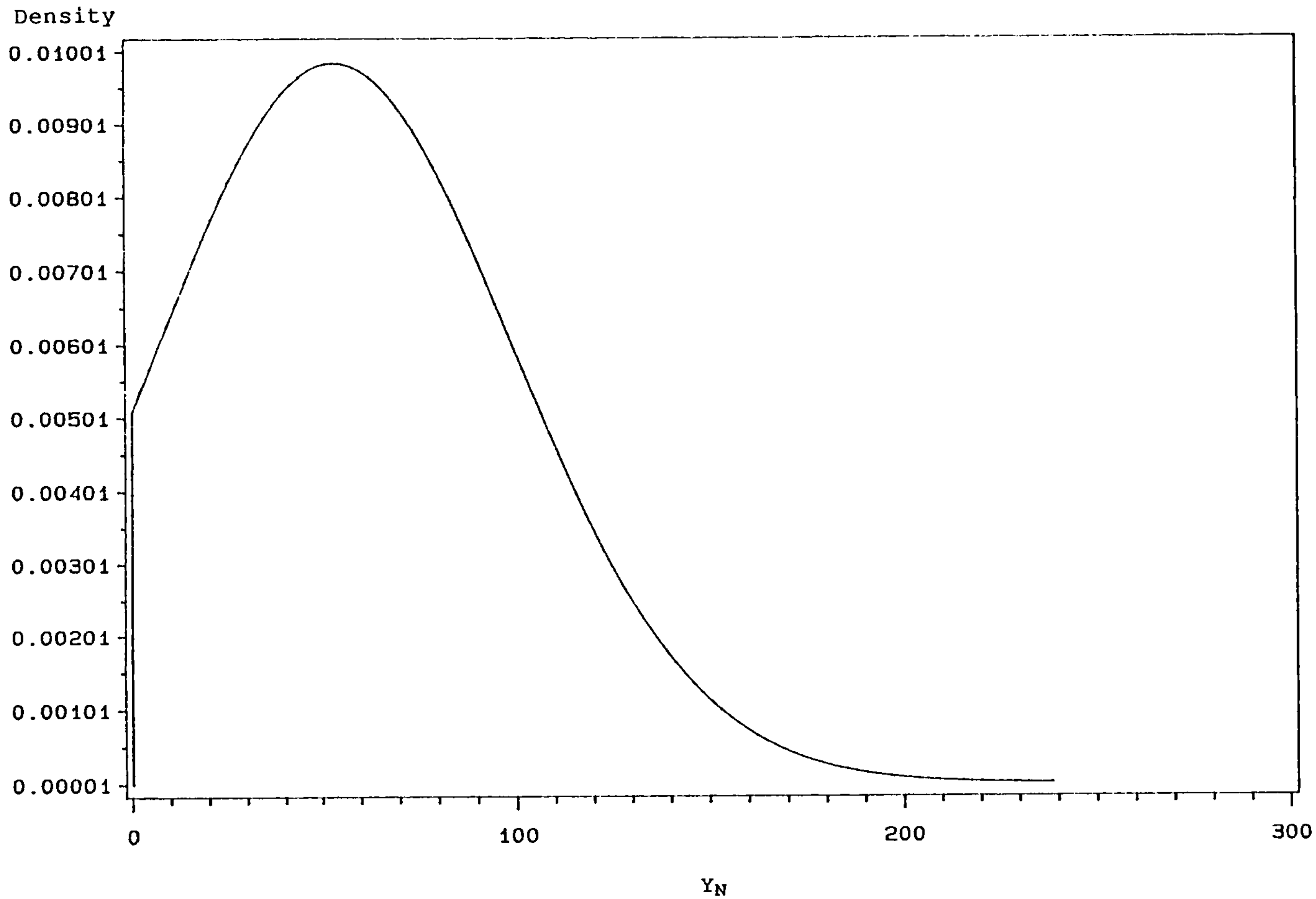

FIGURE III-3 Predictive Distribution for Y_N. (Micrograms U235. Mean 53.1, SD. 46.3, DF 100)

Micrograms U-235

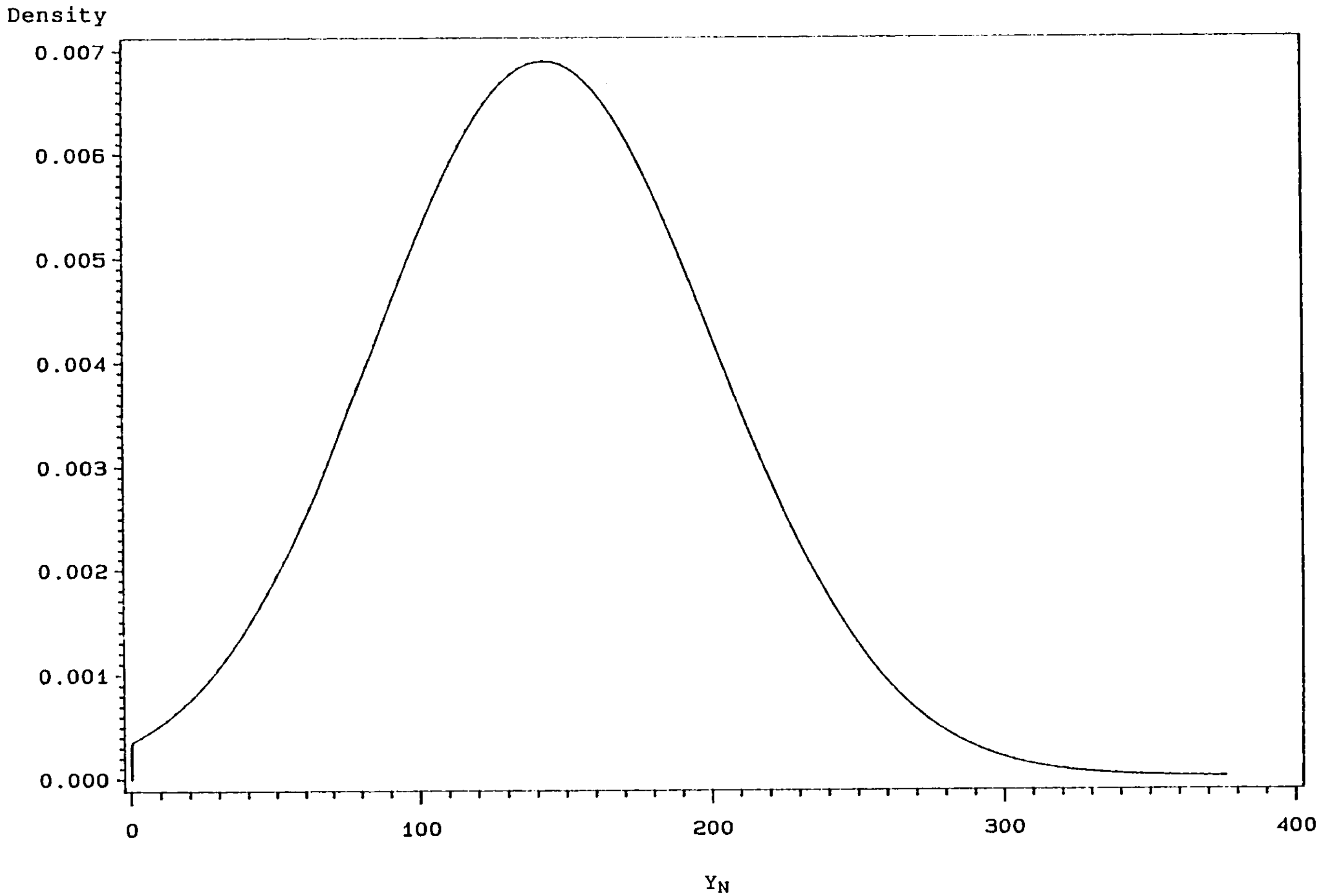

FIGURE III-4 Predictive Distribution for Y_N. (Micrograms U235. Mean 143, SD 58.2, DF 100)

Micrograms U-235

The FORTRAN programs were written for use on an IBM 9370 series minicomputer, using the VM operating system. They could be transferred for use on other computer systems but would require modification of operating system specific sections relating to input and output.

Calibration data pairs for the "CALC" program are supplied in an 80-character fixed record-length file named "FORINP DATA A". One data pair appears on each line; urinalysis measurements in columns 1-10 and lung count measurements in columns 11-20. Decimal points are used in both measurements. Output from the program are written to the file "FOROUT DATA A". No user interaction with this file is required but it is available for inspection.

The program "PREDICT" reads the data stored in file "FOROUT DATA A" to obtain the necessary prediction parameters. It then reads in records from a user-supplied input file, "PRDIN DATA A". Each record in this file has a urinalysis measurement in columns 1-10. Additional identifying information may be included in columns 11-20. This information will be written along with the input urinalysis value and the predictive density distribution to the output file, "PRDOUT DATA A".

CHAPTER IV

EFFECT OF UNCERTAIN URINALYSIS DATA

4.1 Method of Investigation

An assumption underlying the predictive methods employed is that the values of the independent variable, X_i, are fixed and precisely known. This, of course, is not true. While controls existed to minimize error in the urinalysis program at Fernald, random measurement error was inevitable.

As no specific information was available about the measurement uncertainty of the urinalysis program employed at Fernald, it was decided to write a computer program to explore the effect of uncertain input values on the prediction algorithms, initially using the conservative assumption that the variability in the input data was relatively large. Each input measurement was assumed to be the mean of a normal distribution with a standard deviation equal to one-half of the measurement. A random variable was then generated from each distribution to create a new input data set. These data, along with the associated lung counts, were used to calculate the regression parameters and to generate a predictive density distribution for Y_N given a supplied urinalysis measurement, X_N. This process was repeated multiple times to generate a distribution of estimates for $\widehat{B_0}$, $\widehat{B_1}$, and Y_N.

Baselt (1988) indicates that a coefficient of variation as small as 0.1 can be obtained in performing urinalysis by the fluometric method. The simulation procedure was re-run using this value for comparison. Results for these and additional simulations are summarized in Table IV-1.

The results of the simulations follow expected patterns. For instance, as the assumed urinalysis variability increases, the estimate for $\widehat{B_1}$ decreases and the estimate for $\widehat{B_0}$ increases. $\widehat{B_1}$ may be expressed as the sample estimated covariance of the lung counts and the urinalyses divided by the sample estimated variance of the urinalyses. As the variability of the urinalyses increases, this ratio should decrease, as is observed. Likewise, $\widehat{B_0}$ may be expressed as the mean of the lung counts minus the product of $\widehat{B_1}$ and the mean of the urinalyses. As $\widehat{B_1}$ decreases, $\widehat{B_0}$ would be expected to increase, as it does.

4.2 Computer program

A FORTRAN computer program "SIMU" was written to perform the simulation of uncertainty in urinalysis

Table IV-1. Simulation Results

Number of Simulations	Standard Deviation in Micrograms (or fraction of observed value)	Mean $\hat{B}_0$ over simulations	Mean $\hat{B}_1$ over simulations	Mean YN over simulations
5	1.0	46.8	1.14	53.6
10	1.0	45.6	1.16	53.5
100	1.0	46.0	1.20	53.2
1000	1.0	46.0	1.20	53.2
5	(0.1)	47.1	1.12	53.8
10	(0.1)	46.6	1.15	53.5
100	(0.1)	46.0	1.20	53.2
1000	(0.1)	46.1	1.19	53.3
5	(0.5)	54.5	0.45	57.2
10	(0.5)	52.5	0.62	56.2
100	(0.5)	50.5	0.78	55.2
1000	(0.5)	50.9	0.76	55.4

measurements and to calculate regression parameter estimates and predictive distributions based on the simulated measurements (See Appendix C3 for listing).

The program was written to interact with an investigator at runtime, requesting user inputs for:

1. a random number generator seed value,
2. the assumed standard deviation in the urinalysis measurements, and
3. the number of simulation runs to be performed.

The assumed standard deviation may be entered in absolute or relative measure. In other words, the user may request an assumed standard deviation of some fixed number of micrograms per liter or an assumed standard deviation equal to a specific fraction of each measurement. The number of simulation runs may vary between five and a thousand.

The program requires an input calibration data set in the same form as the "CALC" program. It also requires one urinalysis value supplied in the "PRDIN DATA A" file for use in generating example predictive distributions.

After obtaining the user runtime input and reading the input data files, the program generates a random urinalysis value to substitute for each supplied value. Indexing through the input data set, for each supplied urinalysis measurement the program first generates a random number from a uniform distribution. The algorithm used is adapted from one supplied by the netlib@ornl software libary. Using an approach suggested by Hammersley and Handscomb (1964) and an approximation to an inversion of the normal distribution given by Abramowitz and Stegun (1968), this value is then used to generate a random standard normal variate. The random normal variate is multiplied by the assumed standard deviation of the urinalysis measurement and then this quantity is added to the initial measurement which is assumed to represent the mean of the distribution. The new, simulated, urinalysis measurement is then stored separately and the process is repeated for all input urinalysis values, including the value supplied in "PRDIN DATA A" for use in prediction. When this step is complete, the program computes the distributions of B_0, B_1, and Y_N by the same methods as the "CALC" program. At the end of the step, $\widehat{B_0}$, $\widehat{B_1}$, and the mode of Y_N are written to a "PARMS DATA A" file along with the estimates of their standard deviations and the simulation and calculation process is repeated as required. After all simulation runs are complete, the program calculates means over all simulations for these estimates and writes them to the file "SUMARY DATA A". Examples of these files appear in Appendix C4. A number of copies of the SUMARY DATA A file for differerent simulation runs are included. They have been distinguished by naming them SUMARY D1, SUMARY D2, etc.

4.3 Results of Computer Simulation

Simulation runs of 5, 10, 100, and 1000 iterations were performed using the sample data set for calibration and a sample input urinalysis value for prediction. Assumed urinalysis standard deviation values included 1.0, one-tenth of the measurement, and one-half of the measurement. The summary results, presented in Table IV-1., indicate how the regression and prediction parameters respond to relatively large variations in the input urinalysis values, given the assumption that these values were normally distributed. Summary statistics for the runs appear in computer output included in Appendix C4. The distributions of some calculated parameter estimates are presented in Figures IV-1 through IV-6.

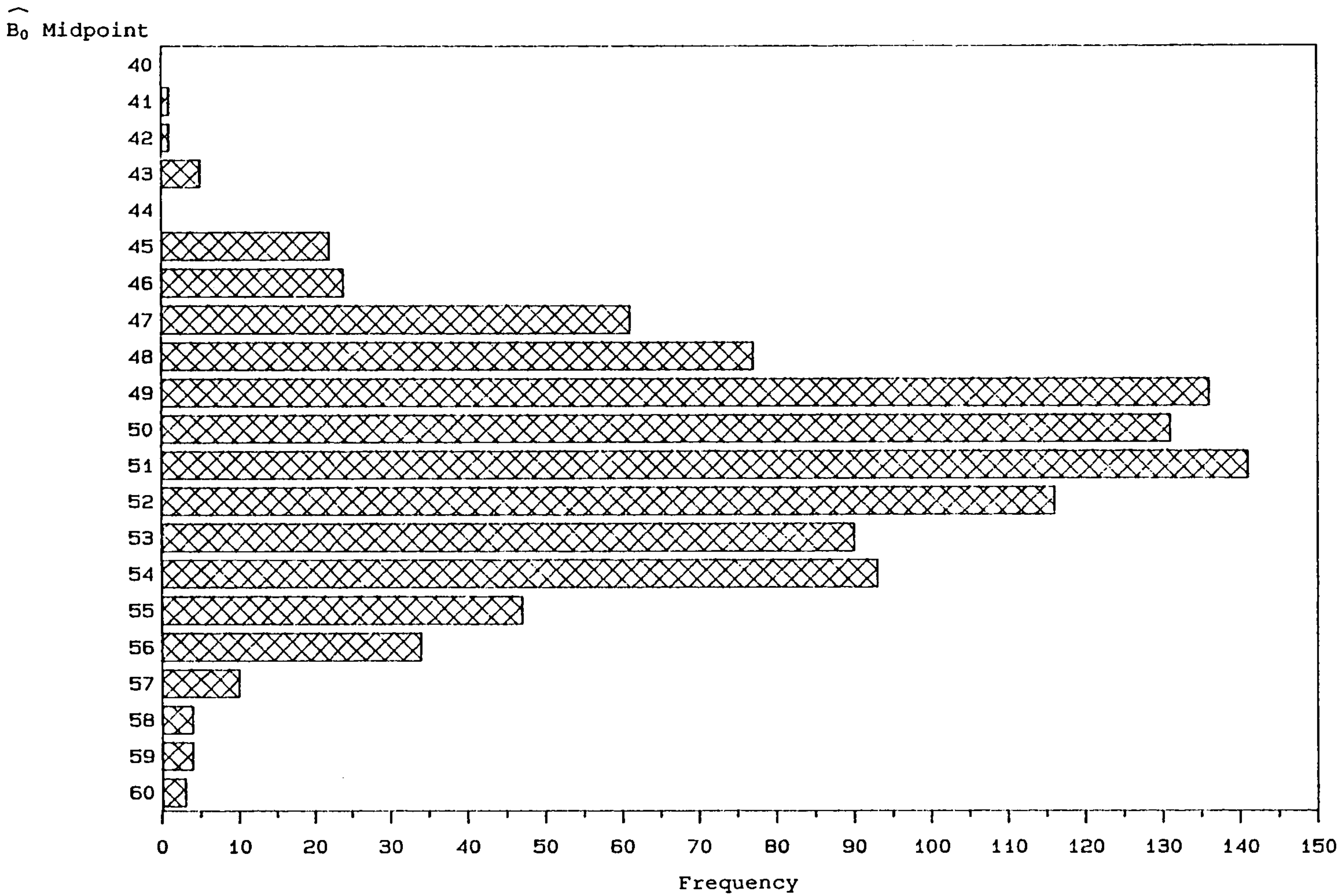

FIGURE IV-1 $\hat{B}_0$ Values Obtained in 1000 Simulations.

Urinalysis SD = 1/2 of Measurement

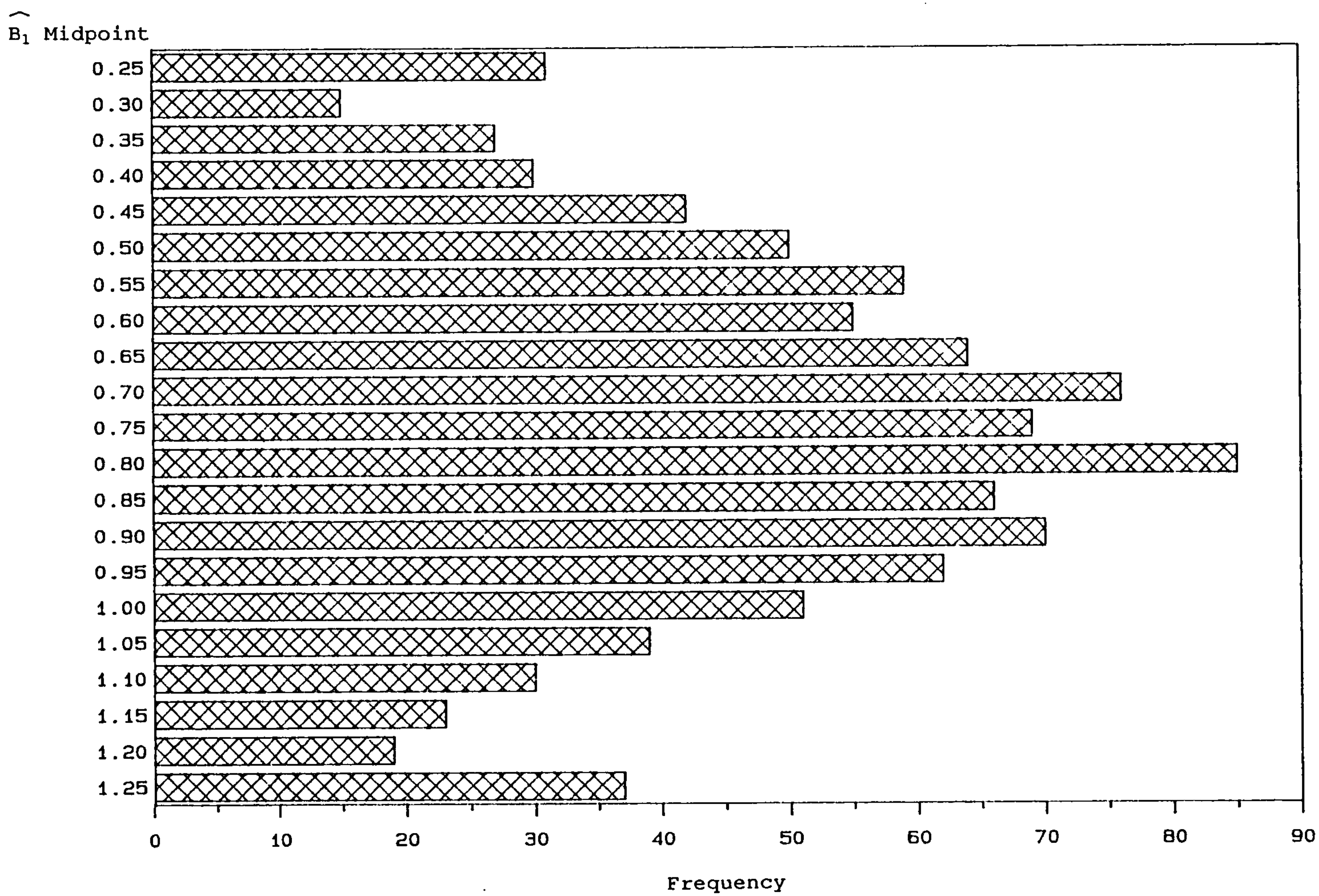

FIGURE IV-2 $\hat{B}_1$ Values Obtained in 1000 Simulations.

Urinalysis SD = 1/2 of Measurement

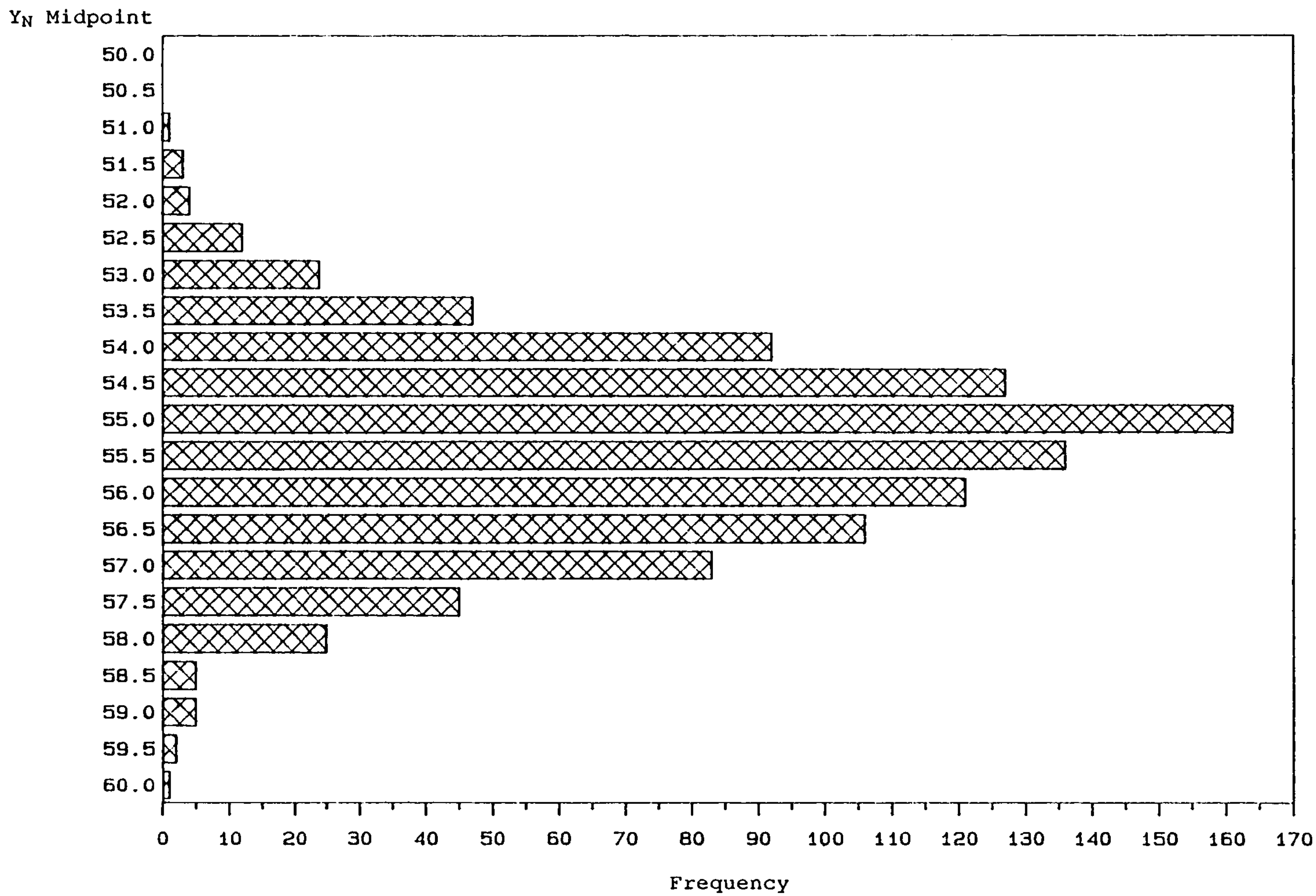

FIGURE IV-3 Y_N Modal Values Obtained in 1000 Simulations.

Urinalysis SD = 1/2 of Measurement

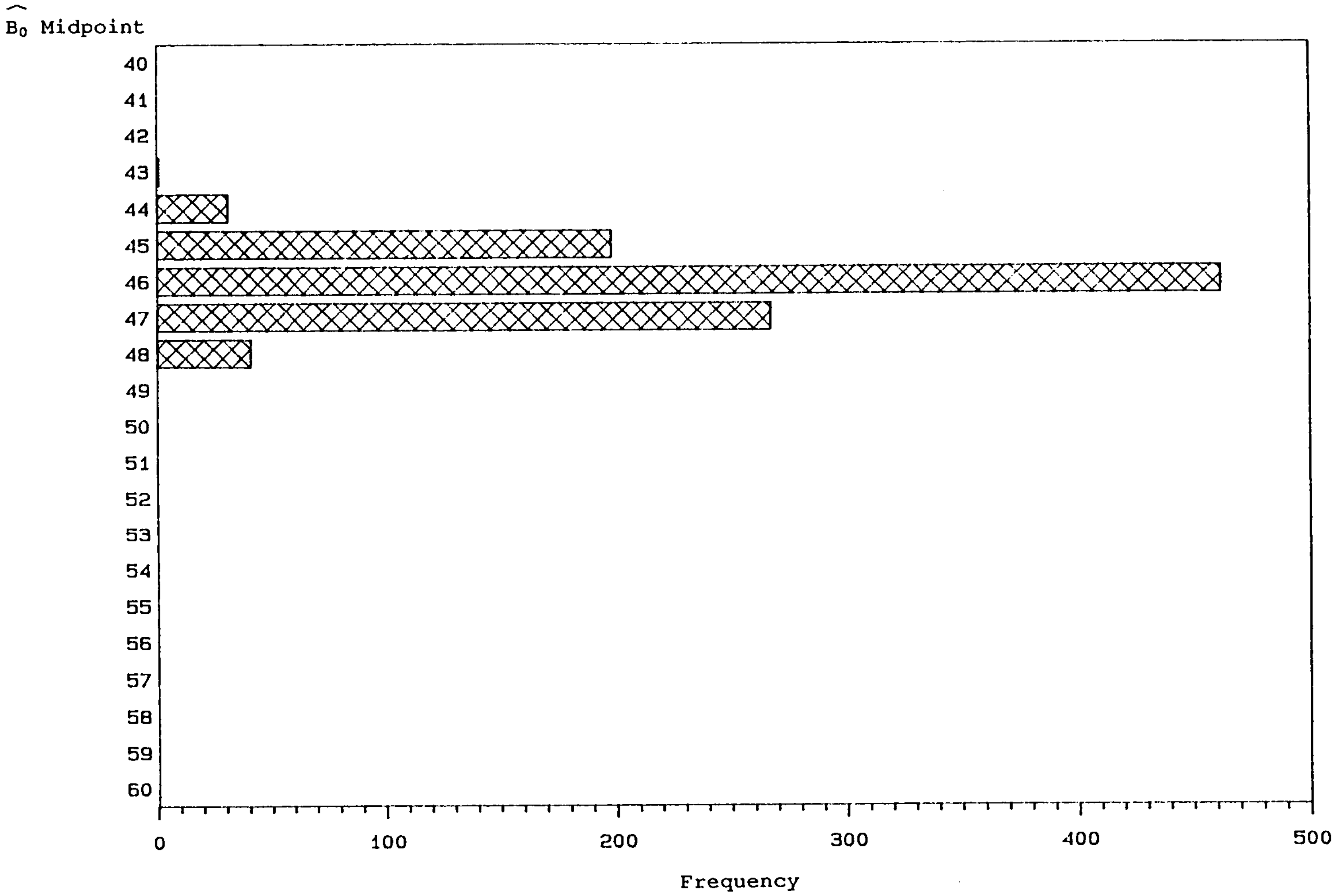

FIGURE IV-4 $\hat{B}_0$ Values Obtained in 1000 Simulations.

Urinalysis SD = 1/10 of Measurement

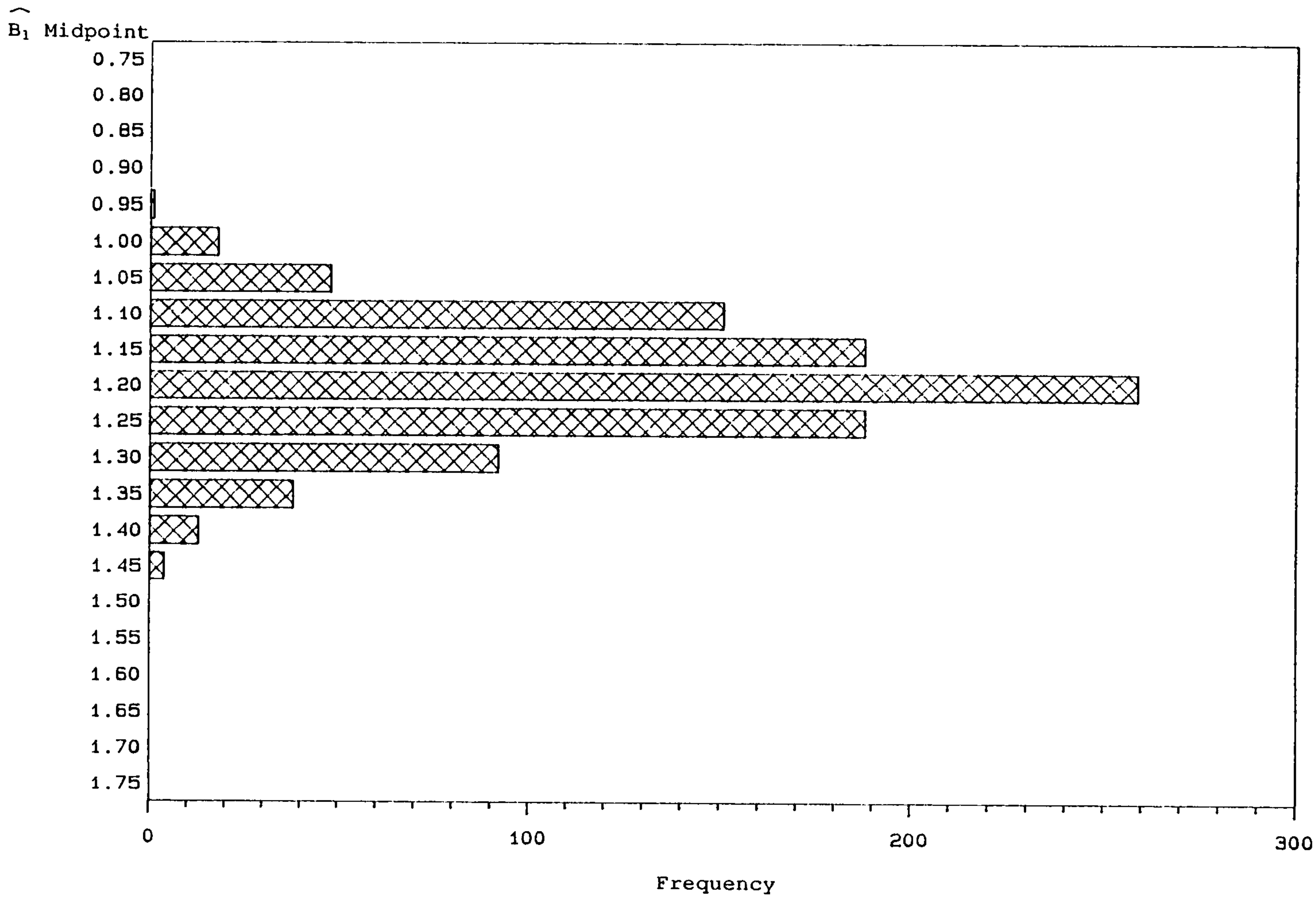

FIGURE IV-5 $\hat{B}_1$ Values Obtained in 1000 Simulations.

Urinalysis SD = 1/10 of Measurement

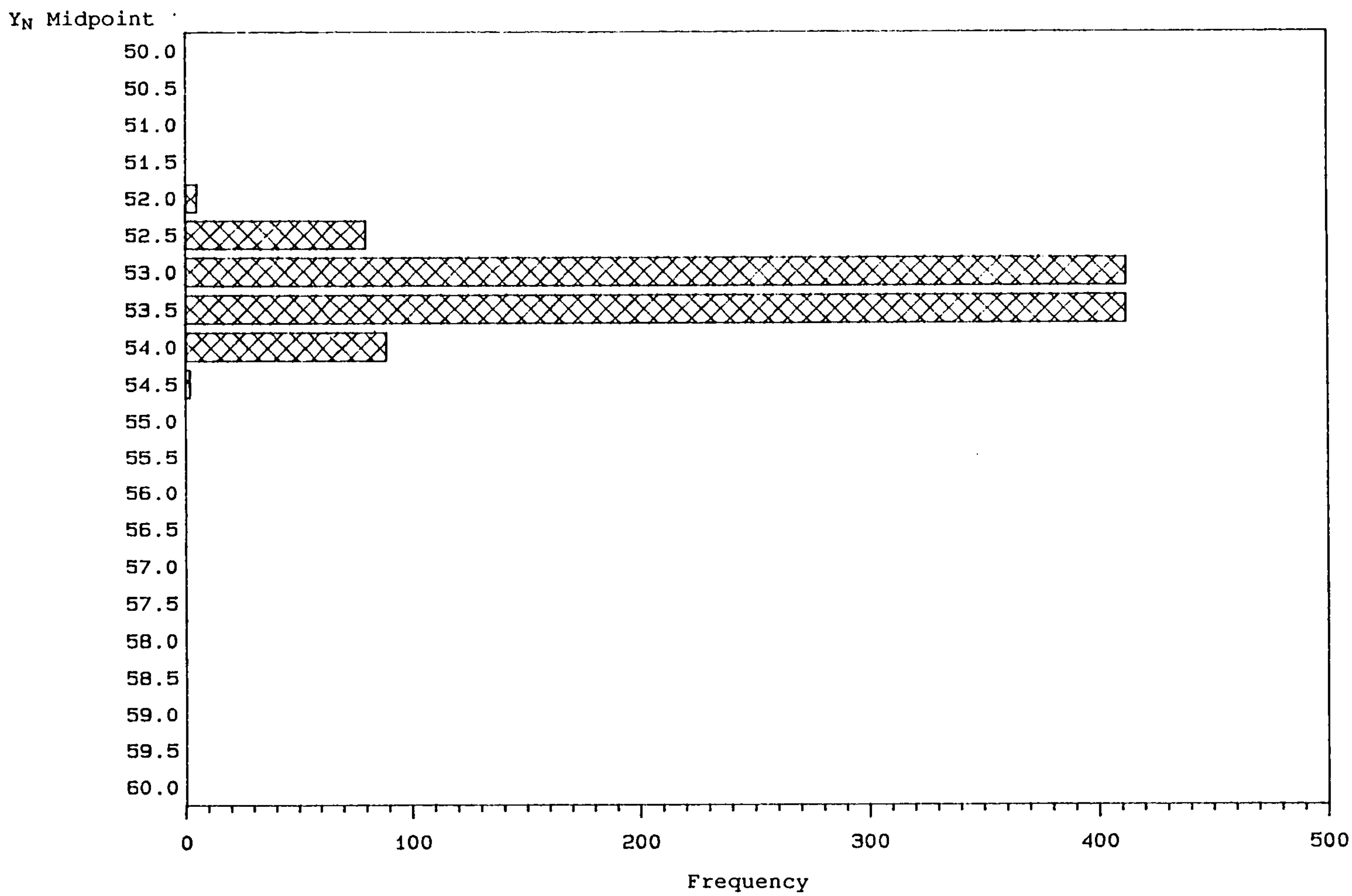

FIGURE IV-6 Y_N Modal Values Obtained in 1000 Simulations.

Urinalysis SD = 1/10 of Measurement

CHAPTER V

DISCUSSION AND SUGGESTIONS FOR FURTHER RESEARCH

5.1 Discussion

Many epidemiologic studies have been proposed or undertaken to investigate the effects of radiation exposures on workers whose doses are not precisely known. The problem described in this thesis, the existence of only limited quantities of lung counting data for groups of uranium workers, has been encountered a number of times. There are sometimes other biological monitoring measures such as urinalysis available, but it is often difficult to estimate lung activity from them.

This research identified a means whereby one of the most common biological monitoring measures, urinalysis, could be used to infer a quantity more useful in epidemiologic studies, lung burden of Uranium-235. By performing a Bayesian linear regression of the relationship between urinalysis and lung count data when both are available it is possible to make predictions of lung activity when only urinalysis is available. The predictions are provided in the form of predictive densities, which indicate the precision of the prediction.

As previously described, this thesis predicted the lung activity due to Uranium-235. The Uranium-238, Uranium-234, and uranium daughter products also contribute to lung dose in uranium workers and must be considered when evaluating dose. Further research is required to identify the best means of estimating the contributions of these isotopes to lung dose.

The statistical model applied to the calibration data in this thesis assumed that the variance of the lung activity measurements did not change with increasing urinalysis values. This simplifying assumption may not be realistic and consideration might be given to relaxing the assumption.

The calibration data set examined herein was selected to be representative of typical radiological monitoring data. The variability in the input data was reflected in the densities of the for model parameters and the predictive distributions for lung burdens. A conclusion which may be drawn here is that inferences about Uranium-235 lung burden and, hence, dose from Uranium-235 may be made from limited urinalysis data but that these inferences are not precise. It is implied that epidemiologic studies employing such measures should take into account this lack of precision in the estimated doses.

5.2 Suggestions for Further Research

The utility of the lung count predictions would be enhanced if the the spread of their probability density functions could be reduced. Bayesian theory suggests a way to do this. Once a posterior density for the parameters of interest has been determined, its precision can be improved by incorporation of additional data. The original posterior density may be used as the prior density in the subsequent analysis and the Bayesian algorithm may be repeated with the new data. This will often result in narrower density functions for the parameters.

In this case, it would be interesting to update the results obtained using data for uranium workers at another facility. Studies have been performed on groups of workers at other facilities and similar data should be available. For example, a large body of radiological monitoring data exists for the Oak Ridge Y-12 weapons plant. If it could be established that the materials handled were chemically similar to those handed at the Fernald FMPC, it would be of interest to attempt to use this data in an update procedure.

Similarly, it would be of interest to use data from another facility with similar exposure conditions to create an informative prior and to re-analyze the FMPC data using this prior.

LIST OF REFERENCES

LIST OF REFERENCES

Abramowitz, M.; Stegun, I.A. Handbook of mathematical functions. New York; Dover Publications, Inc.; 1968.

Anonymous. Mobile in vivo radiation monitoring laboratory. Oak Ridge, TN: Y12 pamphlet Y-DK-86.; 1961.

Baselt, R.C Biological monitoring methods for industrial chemicals, second edition. Littonton, Mass; PSG Publishing; 1988.

Beyer, W.H. CRC standard mathematical tables and formulas. Boca Raton, FL: CRC Press; 1991.

Boback, M.W. A review of uranium excretion and clinical urinalysis data in accidental exposure cases. Proceedings. Conference on Occupational Health Experience with Uranium. Arlington, VA: ERPA 93; 1975.

Boback, M.W.; Heatherton, R.C. Bioassay aspects of a UF6 fume release. Paper presented at the 12th Annual Bioassay and Analytical Chemistry Meeting; Gatlinburg, TN; Oct. 13, 1966.

Chase, Gary R. Correlation between airborne U exposure and U urinalysis results at Bear Creek Uranium. Health Phys. 56:195 - 199; 1989.

Committee on the Biological Effects of Ionizing Radiations (BEIR V), National Research Council, National Academy of Sciences. The effects on populations of exposure to low levels of ionizing radiation. Washington, DC: National Academy Press; 1990.

Fugate, J.R. Information brochure. Uranium Feed Materials Production Center. Records Management Dept., NLO, Inc; 1982.

Gradshteyn, I.S.; Ryzhik, I.M. Table of integrals, series, and products, fourth edition. London; Academic Press; 1965

Heatherton, R.C. Occupational health experience with a contractor uranium refinery. Proceedings. Conference on Occupational Health Experience with Uranium, Arlington, VA; ERDA 93; 1975.

Hammersley, J.M.; Handscomb, D.C. Monte Carlo methods. London; Chapman and Hall; 1979.

Little, R.J.A.; The statistical analysis of low-level radioactivity in the presence of background counts. Health Phys. 43: 693-703; 1982.

Press, S. J. Baysian statistics, principles, models and applications. New York, NY: Wiley and Sons; 1989

Scott, L.M.; West, C.M. Uranium in vivo spectrometry. Proceedings. Conference on Occupational Health Experience with Uranium. Arlington, VA: ERPA 93; 1975.

Smith, J.M.; Moss, C.E., Jr.; Flesch, J.P. Health hazard evaluation report. Washington, DC: US Dept. of Health and Human Services; National Institute of Occupational Safety and Health Report; HETA 85-111-1770; 1987.

Spitz, H.B.; Simpson, J.C.; Aldridge, T.L. Analysis of uranium urinalysis and in vivo measurement results from eleven participating uranium mills. Washington, DC: Batelle Pacific Northwest Laboratory; USNRC (NUREG/CR-2955); 1984.

Walpole, R.E; Myers, R. H. Probability and statistics for engineers and scientists, fourth edition. New York, NY: MacMillan Publishing Company; 1989.

Personal Communication. Memo from C.M. West, Health Physicist, Oak Ridge Associated Universities, Oak Ridge, TN. 1988.

APPENDIXES

APPENDIX A

SAMPLE DATA SET

FILE: FORINP TEMP A VM/IS

DATA SET
PAIRED URINALYSES AND LUNG COUNTS
OCCURRING IN SAME MONTH
PLANT 8 WORKERS
NEGATIVE VALUES SET TO ZERO

URINALYSIS MICROGRAMS PER LITER	LUNG COUNT EST. BURDEN MICROGM. U235
17.00	27.00
7.00	19.00
3.00	0.00
1.00	97.00
10.00	5.00
10.00	0.00
0.00	28.10
8.00	129.00
1.00	112.00
1.00	70.00
21.00	125.00
5.00	79.00
4.00	95.00
20.00	98.00
7.00	102.00
7.00	102.00
10.00	114.00
10.00	33.00
6.00	40.00
1.00	14.00
9.00	58.00
16.00	101.00
8.00	0.00
18.00	116.00
55.00	116.00
4.00	2.80
7.00	30.00
11.00	40.00
18.00	222.00
11.00	78.00
39.00	154.00
4.00	89.00
10.00	32.00
7.00	33.00
3.00	33.00
6.00	33.00
3.00	92.00
28.00	33.00
18.00	0.00
7.00	0.00
11.00	57.00
16.00	93.00
8.00	28.00
11.00	0.00

FILE: FORINP TEMP A VM/IS

10.00	29.00
15.00	17.00
7.00	25.00
1.00	0.00
29.00	59.00
8.00	91.00
18.00	89.00
3.00	67.20
5.00	38.00
18.00	11.00
11.00	11.00
3.00	161.00
1.00	152.00
5.00	0.00
12.00	71.00
18.00	42.00
3.00	110.00
3.00	164.00
9.00	95.00
6.00	97.00
29.00	73.00
14.00	112.00
26.00	52.00
9.00	32.65
7.00	44.10
11.00	17.00
4.00	17.00
5.00	44.00
17.00	24.00
9.00	73.00
8.00	73.00
2.00	27.00
26.00	49.00
19.00	49.00
6.00	49.00
17.00	18.00
7.00	46.00
7.00	72.40
16.00	157.00
16.00	85.00
19.00	102.00
17.00	124.00
23.00	65.00
4.00	23.00
10.00	116.00
7.00	38.00
6.00	2.00
5.00	83.00
2.00	24.00
11.00	24.00
36.00	123.00
7.00	0.00
16.00	0.00
12.00	0.00
7.00	43.00

FILE: FORINP TEMP A VM/IS

```
5.00    58.34
3.00    28.00
5.00     0.00
```

APPENDIX B

DETAILS OF DERIVATIONS

APPENDIX B

ADDITIONAL DETAILS OF DERIVATIONS PRESENTED IN CHAPTER III

Derivations presented here generally follow Press(1989). Different assumptions were made regarding integration limits for B_0 and B_1 while integrating their joint posterior density to determine the marginal densities: $(0 \leq B_0, B_1 < \infty)$

Symbolically, we wish to determine $p(B_0, B_1, \sigma | \underline{X}, \underline{Y})$, the posterior density. This may be obtained by multiplying the likelihood function:

$$p(\underline{Y}, \underline{X} \mid B_0, B_1, \sigma)$$

by the prior density:

$$p(B_0, B_1, \sigma)$$

and dividing by the normalizing factor:

$$\int_0^\infty \int_0^\infty \int_0^\infty p(\underline{Y}, \underline{X} \mid B_0, B_1, \sigma)\ p(B_0, B_1, \sigma)\ dB_0\ dB_i\ d\sigma$$

Subsequently, we will not consider the normalizing factor but will simply express:

$$p(B_0, B_1, \sigma | \underline{X}, \underline{Y}) \propto p(\underline{Y}, \underline{X} \mid B_0, B_1, \sigma) p(B_0, B_1, \sigma).$$

The simple linear regression model is:

$$Y_i \mid X_i = B_0 + B_1 X_i + \mu_i, \quad i=1, \ldots, n$$

B_0 and B_1 are unknown regression parameters. If $\mu_i \ldots \mu_n$ are assumed i.i.d. $N(0, \sigma^2)$, the model parameters are B_0, B_1, σ; and the likelihood of observing X_1, Y_1 is:

$$p\left(Y_1, X_1 \mid B_0, B_1, \sigma\right) \propto \frac{1}{\sigma} e^{\left\{-\frac{1}{2\sigma^2}\left(Y_1 - B_0 - B_1 X_1\right)^2\right\}}$$

The likelihood of observing $\underline{Y} = (Y_1, \ldots Y_n)$ with $\underline{X} = (X_1, \ldots, X_n)$ is:

$$p\left(\underline{Y}, \underline{X} \mid B_0, B_1, \sigma\right) \propto \frac{1}{\sigma} e^{\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - B_0 - B_1 X_i)^2\right\}}$$

Assuming a vague prior density:

$$P(B_0) \propto \text{const.} \qquad P(B_1) \propto \text{const} \qquad P(\sigma) \propto 1/\sigma$$

$$p(B_0, B_1, \sigma) = p(B_0) p(B_1) p(\sigma) \propto 1/\sigma$$

The joint posterior density, then, is:

$$p(B_0,\ B_1,\ \sigma\ |\ \underline{X},\ \underline{Y}) \propto \frac{1}{\sigma} \cdot \frac{1}{\sigma^n}\, e\left\{\frac{1}{-2\sigma^2} \sum_{i=1}^{n} (Y_i - B_0 - B_1 X_i)^2\right\}$$

$$\propto \frac{1}{\sigma^{n+1}}\, e\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (Y_i - B_0 - B_1 X_i)^2\right\}$$

where $B_0 \geq 0$, $B_1 > 0$, $\sigma > 0$. The maximum likelihood estimators for B_0 and B_1 are equivalent to the least squares estimators. The sum of the squared residuals is:

$$\sum_{i=1}^{n} \left(Y_i - B_0 - B_1 X_i\right)^2$$

This expression may be differentiated to obtain the partial derivatives with respect to B_0 and B_1.

$$\frac{\partial}{\partial B_0} \sum_{i=1}^{n} \left(Y_i - B_0 - B_1 X_i\right)^2 = -2 \sum_{i=1}^{n} \left(Y_i - B_0 - B_1 X_i\right)$$

$$\frac{\partial}{\partial B_1} \sum_{i=1}^{n} \left(Y_i - B_0 - B_1 X_i\right)^2 = -2 \sum_{i=1}^{n} \left(Y_i - B_0 - B_1 X_i\right) X_i$$

Setting these expressions to zero and rearranging, the normal equations may be obtained:

$$nB_0 + B_1 \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i$$

$$B_0 \sum_{i=1}^{n} X_i + B_1 \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i$$

and the estimators for B_0 and B_1 may be expressed:

$$\widehat{B_0} = \frac{\sum_{i=1}^{n} Y_i - B_1 \sum_{i=1}^{n} X_i}{n}$$

$$\widehat{B_1} = \frac{n \sum_{i=1}^{n} X_i Y_i - \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{n \sum_{i=1}^{n} X_i^2 - \left(\sum_{i=1}^{n} X_i \right)^2}$$

An unbiased estimator for σ^2 is (Walpole and Myers, 1989):

$$s^2 = \sum_{i=1}^{n} \frac{\left(Y_i - \widehat{B_0} - \widehat{B_1} X_i \right)^2}{n - 2}$$

Noting that:

$$\left(Y_i - B_0 - B_1 X_i \right)^2 =$$

$$\left[\left(Y_i - \widehat{B_0} - \widehat{B_1} X_i \right) - \left(B_0 - \widehat{B_0} \right) - \left(B_1 - \widehat{B_1} \right) X_i \right]^2,$$

the density of the parameters B_0 and B_1 may be obtained by integrating the joint posterior density over σ:

$$\int_0^\infty p\left(B_0,\ B_1,\ \sigma \mid \underline{X},\ \underline{Y}\right) d\sigma \propto \int_0^\infty \frac{1}{\sigma^{n+1}}\ e\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n} (Y_i-B_0-B_1X_i)^2\right\} d\sigma$$

This can be reduced to a standard integral by substituting

$$a = \sum_{i=1}^{n} (Y_i-B_0-B_1X_i)^2$$

and performing a change of variables, letting

$$T = (a/2)\ \sigma^{-2}$$

$$dT = -a\ \sigma^{-3} d\sigma$$

$$d\sigma = -dT/a\sigma^{-3}$$

$$\sigma = \left(T/(a/2)\right)^{-1/2}$$

to obtain

$$-\frac{1}{a}\int_0^{\infty}\left[\left(\frac{T}{a/2}\right)^{-1/2}\right]^{2-n} e^{-T}\, dT = -\frac{1}{a}\cdot\frac{1}{(a/2)^{n/2-1}}\int_0^{\infty} T^{n/2-1}\, e^{-T}\, dT$$

This is equivalent to (Beyer, 1991):

$$\propto -\frac{1}{(a/2)^{n/2}}\,\Gamma\left(\frac{n}{2}\right)$$

$$\propto \quad a^{-n/2}$$

Before substituting

$$\sum_{i=1}^{n}(Y_i-B_0-B_1X_i)^2$$

for a, we rearrange:

$$\sum_{i=1}^{n}(Y_i-B_0-B_1X_i)^2 =$$

$$\sum_{i=1}^{n}\left[\left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right) - \left(B_0-\widehat{B_0}\right) - \left(B_1-\widehat{B_1}\right)X_i\right]^2 =$$

$$\sum_{i=1}^{n}\left[\left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right)^2 + \left(B_0-\widehat{B_0}\right)^2 + \left(B_1-\widehat{B_1}\right)^2 X_i^{\,2}\right]$$

$$- 2\sum_{i=1}^{n}\left[\left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right)\left(B_0-\widehat{B_0}\right) + \left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right)\left(B_1-\widehat{B_1}\right)X_i\right]$$

$$+ 2\sum_{i=1}^{n}\left[\left(B_0-\widehat{B_0}\right)\left(B_1-\widehat{B_1}\right)X_i\right]$$

The second term of this expression may be eliminated if we note that:

$$\sum_{i=1}^{n}\left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right) = 0 \quad \text{and} \quad \sum_{i=1}^{n}\left(Y_i-\widehat{B_0}-\widehat{B_1}X_i\right)X_i = 0$$

Finally, replacing "a" with its equivalent expression and substituting "s" for its definition (see page 29) where possible, we obtain the result:

$$p\left(B_0,\ B_1 \mid \underline{X},\ \underline{Y}\right) \propto$$

$$\left[(n-2)s^2 + n\left(B_0 - \widehat{B_0}\right)^2 + \left(B_1 - \widehat{B_1}\right)^2\sum_{i=1}^{n} X_i^{\,2} + 2\left(B_0 - \widehat{B_0}\right)\left(B_1 - \widehat{B_1}\right)\sum_{i=1}^{n} X_i\right]^{-\frac{n}{2}}$$

B_0 and B_1, given X and Y, are jointly distributed as a bivariate t-distribution. By integrating, the marginal univariate posterior distributions for B_0 and B_1 are obtained:

$$p\left(B_0 \mid \underline{X},\ \underline{Y}\right) =$$

$$\int_0^\infty p\left(B_0,\ B_1 \mid \underline{X},\ \underline{Y}\right) dB_1 \propto$$

$$\int_0^\infty \left[(n-2)s^2 + n\left(B_0-\widehat{B_0}\right)^2 + \left(B_1-\widehat{B_1}\right)^2 \sum_{i=1}^{n} X_i^{\,2} + 2\left(B_0-\widehat{B_0}\right)\left(B_1-\widehat{B_1}\right)\sum_{i=1}^{n} X_i\right]^{-\frac{n}{2}} dB_1$$

If we substitute

$$A = \left[(n-2)s^2 + n\left(B_0 - \widehat{B_0}\right)^2\right]$$

$$B = \sum_{i=1}^{n} X_i^{\,2}$$

$$C = 2\left(B_0 - \widehat{B_0}\right)\sum_{i=1}^{n} X_i$$

The integral may be expressed:

$$\propto \int_0^\infty \left[A - \frac{C^2}{4B} + \left(\sqrt{B}\left(B_1 - \widehat{B_1}\right) + \frac{C}{2\sqrt{B}}\right)^2\right]^{-\frac{n}{2}} dB_1$$

Taking

$$a^2 = A - \frac{C^2}{4B}\ ;\ \text{and}\quad X^2 = \left(\sqrt{B}\left(B_1 - \widehat{B_1}\right) + \frac{C}{2\sqrt{B}}\right)^2$$

This reduces to the standard form

$$\frac{1}{\sqrt{B}} \int_0^{\infty} \left(a^2 + x^2\right)^{-\frac{n}{2}} dx = \frac{1}{\sqrt{B}} \frac{\pi}{2} \frac{(2n - 3)\,!!}{(2n - 2)\,!!} \frac{1}{a^{2n - 1}}$$

(Gradshteyn, 1965) This is proportional to:

$$\frac{1}{a^{2(n - 1)/2}}$$

$$= \left[(n-2)s^2 + n\left(B_1 - \widehat{B_1}\right)^2 - \frac{4\left(B_0 - \widehat{B_0}\right)^2 \left(\sum_{i=1}^{n} X_i\right)^2}{4\sum_{i=1}^{n} X_i^{\,2}} \right]^{-\frac{(n-1)}{2}}$$

$$= \left[1 + \frac{n\sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2}{(n-2)\, s^2 \sum_{i=1}^{n} X_i^{\,2}} \left(B_0 - \widehat{B_0}\right)^2 \right]^{-\frac{(n-1)}{2}}$$

A t-distribution with υ degrees of freedom is distributed:

$$\sim \left(1 + \frac{x^2}{\upsilon}\right)^{-\frac{(\upsilon + 1)}{2}}$$

So, the normalized distribution of B_0 is given:

$$\left[\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{\frac{s^2}{n}\sum_{i=1}^{n}X_i^2}\right]^{1/2} \times (B_0 - \widehat{B_0}) \mid \underline{Y}, \underline{X} \sim t_{n-2}$$

The distribution of B_1 may be evaluated in a similar manner:

$$p\left(B_1 \mid \underline{Y}, \underline{X}\right) \propto$$

$$\int_0^\infty \left[(n-2)s^2 + n\left(B_0-\widehat{B_0}\right)^2 + \left(B_1-\widehat{B_1}\right)^2\sum_{i=1}^{n} X_i + 2\left(B_1-\widehat{B_1}\right)\sum_{i=1}^{n} X_i\left(B_0-\widehat{B_0}\right)\right]^{-\frac{n}{2}} dB_0$$

by letting

$$A = \left[(n-2)s^2 + \left(B_1 - \widehat{B_1}\right)^2\right]$$

$$B = n$$

$$C = 2\left(B_1 - \widehat{B_1}\right)\sum_{i=1}^{N} X_i$$

and then integrating as before, the normalized distribution of B_1 is found:

$$\frac{\frac{B_1 - \widehat{B_1}}{s}}{\left(\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2\right)^{1/2}} \mid \underline{X}, \underline{Y} \sim t_{n-2}$$

Likewise, the predictive density of Y_N may be obtained:

$$p\left(Y_N \mid X_N, \underline{Y}, \underline{X}\right) =$$

$$\int_0^{\infty}\int_0^{\infty}\int_0^{\infty} p\left(Y_N \mid X_N, B_0, B_1, \sigma\right) p\left(B_0, B_1, \sigma \mid \underline{Y}, \underline{X}\right) dB_0\, dB_1\, d\sigma$$

which yields the following distribution for Y_N:

$$\frac{Y_N - B_0 - B_1 X_N}{s\left[1 + \frac{1}{n} + \frac{\left(X_N - \overline{X}\right)^2}{\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}\right]^{\frac{1}{2}}} \Bigg| \underline{X}, \underline{Y} \sim t_{n-2}$$

The posterior density of σ may be obtained by integrating B_0 and B_1 out of the joint posterior density:

$$p\left(\sigma \mid \underline{X}, \underline{Y}\right) \propto \int_0^{\infty}\int_0^{\infty} \frac{1}{\sigma^{n+1}}\, e^{\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(Y_i - B_0 - B_1 X_i\right)^2\right\}} dB_0\, dB_1$$

Substituting

$$\sum_{i=1}^{n} \left(Y_i - B_0 - B_1X_i\right)^2 =$$

$$\left(n-2\right)s^2 + n\left(B_0 - \widehat{B_0}\right)^2 + \left(B_1 - \widehat{B_1}\right)^2 \sum_{i=1}^{n} X_i{}^2 + 2\left(B_0 - \widehat{B_0}\right)\left(B_1 - \widehat{B_1}\right)\sum_{i=1}^{n} X_i$$

and integrating, we obtain

$$p(\sigma \mid \underline{X}, \underline{Y}) \propto \frac{1}{\sigma^{n-1}} e^{-(n-2)\ s^2\ /2\sigma^2}$$

The expectation is given:

$$E(\sigma \mid \underline{X}, \underline{Y}) = \frac{\Gamma\left(\frac{n-3}{2}\right)}{\Gamma\left(\frac{n-2}{2}\right)} \left(\frac{n-2}{2}\right)^{1/2} s$$

and

$$SD\ (\sigma \mid \underline{X}, \underline{Y}) = \left(\frac{\left(n-2\right)s^2}{\left(n-4\right)} - \left[E(\sigma \mid \underline{X}, \underline{Y})\right]^2\right)^{\frac{1}{2}}$$

APPENDIX C1

"CALC" - FORTRAN PROGRAM FOR CALCULATION OF REGRESSION PARAMETER ESTIMATES

FILE: CALC   FORTRAN A   VM/IS

```
      PROGRAM CALC
C
C     MAIN CALCULATION PROGRAM
C
      DOUBLE PRECISION SSQD,BOFACT,B1FACT,SSDEVX,S,ESIG,NN,SIGFAC
      DOUBLE PRECISION SIGMOD
C
      OPEN (UNIT = 2, FILE = 'FOROUT', STATUS = 'NEW')
C
      SIGFAC=0.0
      NN=0.0
      S=0.0
      SSQD=0.
      SSDEVX=0.
      ESIG=0.
      SIGMOD=0.

C     XMN=MEAN X, YMN=MEAN Y, BZERO=INTERCEPT, BONE=SLOPE, SUMX=SUM OF X
C     SUMY=SUM  OF YS, SMRESQ=SUM OF RESIDUALS - SQUARED, SYSQ= SUM OF
C     SQUARED YS, SXSQ=SUM OF SQAUARED XS, SXY=SUM OF PRODS OF X & Y
C     COUNT=NUMBR OF RECS, SSQD=SIGMA SQD EST,SSDEVX= SUM SQD DEVS OF X
C     ESIG=EXPECTAION OF SIGMA, NN=NUMBER OF RECS IN DBLE PRECIS
C     SIGFAC=STD DEV OF STD DEV (SIGMA), SIGMOD=MODE OF SIGMA DISTRIB'N
C
      CALL LR(XMN,YMN,BONE,BZERO,SUMX,SUMY,SYSQ,SXSQ,SXY,SMRESQ,COUNT)
      V=IFIX(COUNT+.000001)
      NN=DBLE(COUNT)
C
      SSQD=SMRESQ/(COUNT-2.0)
      S=DSQRT(SSQD)
      SSDEVX=(SXSQ-(SUMX**2)/COUNT)
      BOFACT=1.0/((SSDEVX/((SSQD/COUNT)*SXSQ))**0.5)
      B1FACT=DSQRT(SSQD)/DSQRT(SSDEVX)
C     CALCS FOR MARGINAL POSTERIOR DENSITY OF SIGMA
      ESIG=(DGAMMA((NN-3.0)/2.0)/DGAMMA((NN-2.0)/2.0))*
     1 DSQRT((NN-2.0)/2.0)*S
      SIGFAC=DSQRT(((NN-2.0)*SSQD/(NN-4.0))-(ESIG*ESIG))
      SIGMOD=S*DSQRT((NN-2)/(NN-1))
C
      WRITE(2,5)
 5    FORMAT(1X,'XMN,YMN,B0,B1')
      WRITE(2,10)XMN,YMN,BZERO,BONE
 10   FORMAT(1X,F14.6,F14.6,F14.6,F14.6)
      WRITE(2,15)
 15   FORMAT(1X,'SUMX,SUMY,SYSQ,SXSQ,SXY')
      WRITE(2,20)SUMX,SUMY,SYSQ,SXSQ,SXY
 20   FORMAT(1X,F14.6,F14.6,F14.6,F14.6,F14.6)
      WRITE(2,25)
 25   FORMAT(1X,'SMRESQ,COUNT,SSDEVX,S')
      WRITE(2,30)SMRESQ,COUNT,SSDEVX,S
 30   FORMAT(1X,F14.6,F14.6,F14.6,F14.6)
      WRITE(2,35)
 35   FORMAT(1X,'BOFACT,B1FACT')
      WRITE(2,40)BOFACT,B1FACT
 40   FORMAT(1X,F14.6,F14.6)
```

```
      WRITE(2,45)
 45   FORMAT(1X,'ESIG,SIGFAC')
      WRITE(2,50)ESIG,SIGFAC
 50   FORMAT(1X,F14.6,F14.6)
      WRITE(2,55)
 55   FORMAT(1X,'SIGMOD')
      WRITE(2,60)SIGMOD
 60   FORMAT(1X,F14.6)
      CLOSE (2)
      STOP
      END
C
      SUBROUTINE LR(XMN,YMN,BONE,BZERO,SUMX,SUMY,SYSQ,
     1 SXSQ,SXY,SMRESQ,COUNT)
C
C     DOES LIN REGR ON DATA IN FORINP FILE SPECIFIED IN FORTFIX EXEC
C     RETURNS LR PARMS AND OTHER STAT QUANTITIES
C
      REAL X,Y,Z,SUMX,SUMY,SYSQ,SXSQ,SXY,COUNT,BZERO,BONE,YPRED
      REAL RESD,RESDSQ,SMRESD,SMRESQ,SMSQMN,SMSQAR,SMSQDR,YMN,XMN
C
      OPEN (UNIT = 1, FILE = 'FORINP', STATUS = 'OLD')
C
      XMN=0.
      YMN=0.
      RESD=0.
      RESDSQ=0.
      SMRESD=0.
      SMRESQ=0.
      SMSQMN=0.
      SMSQRG=0.
      SMSQAR=0.
      SMSQDR=0.
      BONE=0.
      BZERO=0.
      YI=0.
      COUNT=0.
      SUMX= 0.
      SUMY= 0.
      SYSQ= 0.
      SXSQ= 0.
      SXY= 0.
C
5     READ(1,10,END=500) X,Y
10    FORMAT(F10.4,2X,F10.4)
C
```

FILE: CALC FORTRAN A VM/IS

```
      SUMX=SUMX+X
      SUMY=SUMY+Y
      SXSQ=SXSQ+X*X
      SYSQ=SYSQ+Y*Y
      SXY=SXY+X*Y
      COUNT=COUNT+1.
      GOTO 5
C
500   XMN=SUMX/COUNT
      YMN=SUMY/COUNT
      BONE=(SXY-(SUMX*SUMY/COUNT))/(SXSQ-(SUMX*SUMX/COUNT))
      BZERO=(SUMY/COUNT)-BONE*(SUMX/COUNT)
      CLOSE (1)
      OPEN(UNIT=1,FILE='FORINP',STATUS='OLD')
600   READ(1,610,END=900) X,Y
610   FORMAT(F10.4,2X,F10.4)
      YPRED=BZERO+BONE*X
      RESD=Y-YPRED
      RESDSQ=RESD*RESD
      SMRESD=SMRESD+RESD
      SMRESQ=SMRESQ+RESDSQ
      SMSQMN=SMSQMN+(Y-YMN)*(Y-YMN)
      SMSQAR=SMSQAR+(Y-YPRED)*(Y-YPRED)
      SMSQDR=SMSQDR+(YPRED-YMN)*(YPRED-YMN)
      GOTO 600
900   CLOSE (1)
      RETURN
      END
```

APPENDIX C2

"PREDICT" – FORTRAN PROGRAM FOR CALCULATION OF PREDICTIVE DENSITY

FILE: PREDICT FORTRAN A VM/IS

```
      PROGRAM PREDIC
C
C     MAIN CONTROL PROGRAM FOR PREDICTION OF Y'S  (PERHAPS U235
C     LUNG COUNTS) BASED ON X'S (PERHAPS URINALYSIS MEASUREMENTS
C     OF URANIUM - MASS UNITS - )
C
C
      DOUBLE PRECISION YNEW,XNEW,BZERO,BONE,SMRESQ,COUNT,XMN,SSDEVX
      CHARACTER * 10 ID
C
      OPEN (UNIT = 2, FILE = 'FOROUT', STATUS = 'OLD')
C
      YNEW=0.0
      XNEW=0.0
      BZERO=0.0
      BONE=0.0
      SMRESQ=0.0
      COUNT=0.0
      XMN=0.0
      SSDEVX=0.
C
C
C     XNEW= NEW X, YNEW=PRED Y, BZERO=INTERCEPT, BONE=SLOPE,
C     SMRESQ=SUM OF RESIDUALS - SQUARED, COUNT= NUMBR OF OBS
C     XMN=MEAN X, SSDEVX= SUM OF SQUARED DEVIATIONS OF X FROM MEAN
C     ID=IDENTIFIER FOR RECORD
C
      READ(2,10) DUMMY
 10   FORMAT(1A1)
C     SKIPPING A RECORD
      READ(2,20) XMN,BZERO,BONE
 20   FORMAT(1X,F14.6,14X,F14.6,F14.6)
      READ(2,30) DUMMY
 30   FORMAT(1A1)
      READ(2,32) DUMMY
 32   FORMAT(1A1)
      READ(2,34) DUMMY
 34   FORMAT(1A1)
      READ(2,40) SMRESQ,COUNT,SSDEVX
 40   FORMAT(1X,F14.6,F14.6,F14.6)
      CALL PR(BZERO,BONE,SMRESQ,COUNT,XMN,SSDEVX)
      CLOSE (2)
      STOP
      END
C
C
C
      SUBROUTINE PR(BZERO,BONE,SMRESQ,COUNT,XMN,SSDEVX)
C
      DOUBLE PRECISION BZERO,BONE,SMRESQ,COUNT,XMN,SSDEVX
      DOUBLE PRECISION S,YNFACT,YNEW,XNEW,DF
      CHARACTER * 10 ID
C
      OPEN (UNIT = 3, FILE = 'PRDIN', STATUS = 'OLD')
      OPEN (UNIT = 4, FILE = 'PRDOUT', STATUS = 'NEW')
```

```
C
      XNEW=0.0
      YNEW=0.0
      YNFACT=0.0
      S=0.0
      DF=0.0
C
C
 5    READ(3,10,END=500) XNEW,ID
 10   FORMAT(F14.6,10A)
      S=DSQRT(SMRESQ/(COUNT-2.0))
      YNEW=BZERO+BONE*XNEW
      YNFACT=S*DSQRT(1+(1.0/COUNT)+(((XNEW-XMN)**2)/SSDEVX))
      DF=COUNT-2.0
      WRITE(4,20)XNEW,YNEW,DF,YNFACT,ID
 20   FORMAT(1X,F14.6,1X,F14.6,1X,F14.6,1X,F14.6,10A)

      GOTO 5

500   CLOSE (3)
      CLOSE (4)
       RETURN
       END
```

APPENDIX C3

"SIMU" - FORTRAN PROGRAM TO SIMULATE EFFECTS OF VARIABLITY IN URINALYSIS DATA

```
FILE: SIMU     FORTRAN  A   VM/IS

      PROGRAM SIMU
C
C     MAIN CONTROL PROGRAM
C*********************************************************************
C*  THIS PROGRAM EXPLORES THE EFFECTS OF UNCERTAIN INPUT URINALYSIS  *
C*  DATA ON BAYESIAN LINEAR REGRESSION AND PREDICTION METHODS.       *
C*  THE PROGRAM OPERATES ON A SUPPLIED SET OF URINALYSIS AND LUNG    *
C*  COUNT MEASUREMENTS (FILE: FORINP DATA A).  EACH URINALYSIS       *
C*  MEASUREMENT IS ASSUMED TO COME FROM A NORMAL DISTRIBUTION WITH   *
C*  MEAN=VALUE REPORTED AND STD DEV= (USER RUNTIME INPUT).           *
C*  RUNTIME INPUT).                                                  *
C*  A NORMAL RANDOM VARIABLE IS GENERATED FROM THIS DISTRIBUTION     *
C*  AND THE PROCESS IS  REPEATED FOR ALL INPUT URINALYSIS VALUES.    *
C*  A LINEAR REGRESSION OF THE LUNG COUNTS VS. THE SIMULATED         *
C*  URINALYSES IS PERFORMED AND THE SLOPE (BONE) AND INTERCEPT       *
C*  (BZERO) PARAMETERS ARE STORED.  A LUNG COUNT PREDICTION IS MADE  *
C*  FOR A REPRESENTATIVE URINALYSIS VALUE AND THE RESULT IS STORED.  *
C*  THE STANDARD DEVIATION FOR EACH IS ALSO CALCULATED AND STORED.   *
C*  THE VALUES ARE OUTPUT TO A FILE (PARMS DATA A) AND THE PROCESS   *
C*  IS REPEATED N=-USER RUNTIME INPUT- TIMES.  SUMMARY STATISTICS    *
C*  ARE GENERATED FOR THE INPUT DATA AND CALCULATED PARAMETERS AND   *
C*  THESE ARE WRITTEN TO A SUMMARY FILE (SUMARY DATA A).             *
C*  MEAN OF EACH PARAMETER IS CALCULATED=SUM(PARAMETER)/N            *
C*  STD DEV IS CALCULATED=SUM OVER I (PARAMETER(I)-MEAN)**2 / N-1    *
C*                                                                   *
C*  B DOLAN  FALL 1991                                               *
C*********************************************************************
      INTEGER IY,I,J
      REAL INDAT,OUTDAT,PRIN,PROUT,IFL,MNUR,MNLC
      REAL SUR,SLC,SLCSQ,SURSQ,SURXLC,NFL,SB0,SSQB0,SB0F,SSQB0F,SB1
      REAL SSQB1,SB1F,SYNF,SSQYNF,SPR,SSQSPR,MNB0,MNB1,MNB0F,MNB1F
      REAL MNYNF,MNPR,SDB0,SDB0F,SDB1,SDB1F,SDYNF,SDPR
      REAL SSQD,B0FACT,B1FACT,SSDEVX,S,ESIG,SIGFAC
      REAL MNESIG,SDESIG,SESIG,SSQES,MNSIGF,SDSIGF,SSIGF,SSQSGF
      REAL T1,T2,T3,T4,SDES,SDEF,MNES,MNEF
C
      DIMENSION INDAT(3,1000)
      DIMENSION OUTDAT(8,1000)
C
C     INDAT(1,I)=URINALYSIS(I)            ACTUAL MG/L FROM DATA PAIR
C     INDAT(2,I)=LUNG COUNT(I)            ACT UG U235 FROM DATA PAIR
C     INDAT(3,I)=SIMULATED URINALYSIS(I)  SIMULATED URINALYIS MG/L
C
C     OUTDAT(1,J)=B0(J)                   MAX LIKL EST B0 | DATA
C     OUTDAT(2,J)=B0FACT(J)               CALC STD DEV B0
C     OUTDAT(3,J)=B1(J)                   MAX LIKL EST B1 | DATA
C     OUTDAT(4,J)=B1FACT(J)               CALC STD DEV B1
C     OUTDAT(5,J)=YNFACT(J)               CALC STD DEV PREDICTION
C     OUTDAT(6,J)=PRED(J)                 MEAN OF PREDICTIVE DIST
C     OUTDAT(7,J)=ESIG(J)                 EXPECTATION OF SIGMA | DATA
C     OUTDAT(8,J)=SIGFAC(J)               CALC STD DEV OF SIGMA
C
C     I       INDEX FOR INPUT DATA SET
C     IY      SEED FOR NEXT CALCULATION OF RANDOM NUMBER
C     J       INDEX FOR OUTPUT DATA SET
```

```
C      PRIN   INPUT VALUE FOR PREDICTION
C      PROUT  OUPUT PREDICTED VALUE
C      USDEV  ASSUMED STANDARD DEVIATION OF URINALYSIS VALS
C      N      NUMBER OF SIMULATIONS, <=1000
C      SUR    SUM OF URINALYSES
C      SLC    SUM OF LUNG COUNTS
C      SLCSQ  SUM OF SQUARED LUNG COUNTS
C      SURSQ  SUM OF SQUARED URINALYSES
C      SURXLC SUM OF PROD OF U(I) * LC(I)
C      NFL=N, # SIMULATION RUNS, FLOATED TO A REAL #
C      SBO    SUM OF BO'S
C      SSQBO  SUM OF SQUARED BO'S
C      SBOF   SUM OF BO "FACTORS" (STD DEV)
C      SSQBOF SUM OF SQUARED BO SD
C      SB1    SUM OF B1'S
C      SSQB1  SUM OF SQUARED B1'S
C      SB1F   SUM OF B1 STD DEV'S
C      SSQB1F SUM OF SQUARED B1 STD DEV'S
C      SYNF   SUM OF YNEW STD DEV'S
C      SSQYNF SUM OF SQUARED YNEW STD DEV'S
C      SPR    SUM OF PREDICTED YNEWS
C      SPSQPR SUM OF SQUARED PREDICTED YNEWS
C      MNBO   MEAN BO
C      MNB1   MEAN B1
C      MNBOF  MEAN BOF
C      MNB1F  MEAN B1F
C      MNYNF  MEAN YNF
C      MNPR   MEAN OF PREDICTED VALS
C      MNES   MEAN OF CALCULATED EXPECTATIONS OF SIGMA
C      MNEF   MEAN OF CALCULATED STD DEVIATIONS OF SIGMA
C      SDBO   STD DEV BO
C      SDB1   STD DEV B1
C      SDBOF  STD DEV BOF
C      SDB1F  STD DEV B1F
C      SDYNF  STD DEV YNEWF
C      SDPR   STD DEV PREDICTED YNEWS
C      S      EST STD DEV OF ERRORS
C      ESIG   EXPECTATION OF SIGMA
C      SIGFAC STD DEV OF SIGMA
C      MNESIG MEAN OF CALC SIGMA EXPECTATIONS
C      SDESIG STD DEV OF CALC SIGMA EXPECTATIONS
C      SESIG  SUM OF SIGMA EXPECTATIONS
C      SSQES  SUM OF SQUARED SIGMA EXPECTATIONS
C      MNSIGF MEAN OF CALC SIGMA STD DEVIATIONS
C      SDSIGF STD DEV OF CALC SIGMA STD DEVIATIONS
C      SSIGF  SUM OF CALC SIGMA STD DEVIATIONS
C      SSQSGF SUM OF SQUARED CALC SIGMA STD DEVIATIONS
C      SDES   STD DEV OF SIGMA EXPECTATIONS
C      SDEF   STD DEV OF SIGMA STANDARD DEVIATIONS
C      T1..T4 TEMP VARIABLES
C
C
       OPEN (UNIT = 1, FILE = 'FORINP', STATUS = 'OLD')
       OPEN (UNIT = 3, FILE = 'PROIN' , STATUS = 'OLD')
       OPEN (UNIT = 2, FILE = 'PARMS' , STATUS = 'NEW')
```

```
      OPEN (UNIT = 4, FILE = 'SUMARY', STATUS = 'NEW')
C
C
      PROUT=0.
      PRIN=0.
      S=0.0
      I=0
      J=0
      USDEV=1.0
      N=1
      SSQB0=-0.
      SB1=0.
      SUR=0.
      SLCSQ=0.
      SSQ0F=0.
      SSQB1=0.
      SLC=0.
      SURSQ=0.
      SSQB0F=0.
      SB1F=0.
      SYNF=0.
      SSQYNF=0.
      SPR=0.
      SSQPR=0.
      MNB0=0.
      MNB1=0.
      MNB0F=0.
      MNB1F=0.
      MNYNF=0.
      MNPR=0.
      SDB0=0.
      SDB0F=0.
      SDB1=0.
      SDB1F=0.
      SDYNF=0.
      SDPR=0.
      SB0=0.
      SDUR=0.
      SDLC=0.
      SB0F=0.
      SSQB1F=0.
      ESIG=0.
      SIGFAC=0.
      MNESIG=0.
      SDESIG=0.
      SESIG=0.
      SSQES=0.
      MNSIGF=0.
      SDSIGF=0.
      SSIGF=0.
      SSQSGF=0.
      MNES=0.
      MNEF=0.
      SDES=0.
      SDEF=0.
```

```
      T1=0.
      T2=0.
      T3=0.
      T4=0.
C
      WRITE(6,1)
 1    FORMAT(54H INPUT INTEGER SEED FOR RANDOM NUMBER GENERATOR (OON):)
      READ(5,2)IY
 2    FORMAT(I3)
      WRITE(6,3)
 3    FORMAT(54H INPUT EST. URINALAYSIS MEAS STD DEV-INCLUDE DECIMAL: )
      WRITE(6,4)
 4    FORMAT(52H (USE "-" SIGN IF STDDEV IS FRACT OF URINALYSIS VAL))
      READ(5,5)USDEV
 5    FORMAT(F8.3)
      WRITE(6,6)
 6    FORMAT(54H INPUT INTEGER NUMBER SIMULATION RUNS DESIRED (OOOON):)
      READ(5,7)N
 7    FORMAT(I5)
C*********************************************************************
C RANDOM NUMBER SEED MUST BE 3-DIGIT INTEGER                         *
C STD DEV MUST BE REAL # WITH DECIMAL. MINUS SIGN MAY BE USED TO     *
C INDICATE THAT INPUT VALUE IS A FRACTION TO MULTIPLY EACH MEASUREMENT *
C BY TO ESTIMATE STD DEV.                                            *
C NUMBER OF SIMULATIONS, 5-1000 MUST BE ENTERED AS INTEGER: 00000    *
C*********************************************************************
C
      I=1
 13   READ(1,14,END=15)INDAT(1,I),INDAT(2,I)
 14   FORMAT(F10.3,F10.3)
      I=I+1
      GOTO 13
 15   CONTINUE
      I=I-1
C
      READ(3,17)PRIN
 17   FORMAT(F10.3)
      DO 99 J=1,N
      CALL DATSIM(IY,INDAT,I,PRIN,PROUT,USDEV)
      CALL LRPR(INDAT,OUTDAT,J,I,N,PRIN,PROUT)
 99   CONTINUE
C
      MNLC=0.
      MNUR=0.
      SUR=0.
      SLC=0.
      SLC=0.
      SLCSQ=0.
      SURSQ=0.
      SURXLC=0.

      DO 105 K=1,I
      IFL=FLOAT(I)
      MNUR=MNUR+INDAT(1,K)/IFL
      MNLC=MNLC+INDAT(2,K)/IFL
```

```
      SUR=SUR+INDAT(1,K)
      SLC=SLC+INDAT(2,K)
      SLCSQ=SLCSQ+INDAT(2,K)*INDAT(2,K)
      SURSQ=SURSQ+INDAT(1,K)*INDAT(1,K)
      SURXLC=SURXLC+INDAT(1,K)*INDAT(2,K)
  105 CONTINUE
C
      SDUR=SQRT((SURSQ-((SUR*SUR)/IFL))/(IFL-1.0))
      SDLC=SQRT((SLCSQ-((SLC*SLC)/IFL))/(IFL-1.0))
C
      DO 110 K=1,N
      NFL=FLOAT(N)
      SBO=SBO+OUTDAT(1,K)
      SSQBO=SSQBO+OUTDAT(1,K)*OUTDAT(1,K)
      SBOF=SBOF+OUTDAT(2,K)
      SSQBOF=SSQBOF+OUTDAT(2,K)*OUTDAT(2,K)
      SB1=SB1+OUTDAT(3,K)
      SSQB1=SSQB1+OUTDAT(3,K)*OUTDAT(3,K)
      SB1F=SB1F+OUTDAT(4,K)
      SSQB1F=SSQB1F+OUTDAT(4,K)*OUTDAT(4,K)
      SYNF=SYNF+OUTDAT(5,K)
      SSQYNF=SSQYNF+OUTDAT(5,K)*OUTDAT(5,K)
      SPR=SPR+OUTDAT(6,K)
      SSQPR=SSQPR+OUTDAT(6,K)*OUTDAT(6,K)
      SESIG=SESIG+OUTDAT(7,K)
      SSQES=SSQES+OUTDAT(7,K)*OUTDAT(7,K)
      SSIGF=SSIGF+OUTDAT(8,K)
      SSQSGF=SSQSGF+OUTDAT(8,K)*OUTDAT(8,K)
  110 CONTINUE
C
      MNBO=SBO/NFL
      MNB1=SB1/NFL
      MNBOF=SBOF/NFL
      MNB1F=SB1F/NFL
      MNYNF=SYNF/NFL
      MNPR=SPR/NFL
      MNES=SESIG/NFL
      MNEF=SSIGF/NFL
      SDBO=SQRT((SSQBO-(SBO*SBO)/NFL)/(NFL-1.0))
      SDBOF=SQRT((SSQBOF-(SBOF*SBOF)/NFL)/(NFL-1.0))
      SDB1=SQRT((SSQB1-(SB1*SB1)/NFL)/(NFL-1.0))
      SDB1F=SQRT((SSQB1F-(SB1F*SB1F)/NFL)/(NFL-1.0))
C
C     /* HANDLE NUMERICAL PROBS W/ DATA HAVING NO VARIATION  */
C     *
      T1=(SSQYNF-(SYNF*SYNF)/NFL)/(NFL-1.0)
      IF (T1 .GT. 0.) GOTO 190
      T1=.000001
  190 SDYNF=SQRT(T1)
      T2=(SSQPR-(SPR*SPR)/NFL)/(NFL-1.0)
      IF (T2 .GT. 0.) GOTO 192
      T2=.000001
  192 SDPR=SQRT(T2)
C
      T3=(SSQES-(SESIG*SESIG)/NFL)/(NFL-1.0)
```

```
      IF (T3 .GT. 0.) GOTO 194
      T3=.000001
 194  SDES=SQRT(T3)
      T4=(SSQSGF-(SSIGF*SSIGF)/NFL)/(NFL-1.0)
      IF (T4 .GT. 0.) GOTO 196
      T4=.000001
 196  SDEF=SQRT(T4)
C
      WRITE(4,201)MNBO
 201  FORMAT(F10.6,15H   MEAN BO      )
      WRITE(4,203)SDBO
 203  FORMAT(F10.6,15H   STD DEV BO   )
      WRITE(4,205)MNBOF
 205  FORMAT(F10.6,15H   MEAN BOF     )
      WRITE(4,207)SDBOF
 207  FORMAT(F10.6,15H   STD DEV BOF )
      WRITE(4,209)MNB1
 209  FORMAT(F10.6,15H   MEAN B1      )
      WRITE(4,212)SDB1
 210  FORMAT(F10.6,15H   STD DEV B1   )
      WRITE(4,211)MNB1F
 211  FORMAT(F10.6,15H   MEAN B1F     )
      WRITE(4,212)SDB1F
 212  FORMAT(F10.6,15H   STD DEV B1F )
      WRITE(4,213)MNYNF
 213  FORMAT(F10.6,15H   MEAN YNEWF   )
      WRITE(4,215)SDYNF
 215  FORMAT(F10.6,15H   STD DEV YNF )
      WRITE(4,217)MNPR
 217  FORMAT(F10.6,15H   MEAN PRED    )
      WRITE(4,219)SDPR
 219  FORMAT(F10.6,15H   STD DEV PRED)
      WRITE(4,221)I
 221  FORMAT(I8,15H      I, # DATA )
      WRITE(4,223)N
 223  FORMAT(I8,15H      N, # SIMU )
      WRITE(4,225)MNUR
 225  FORMAT(F10.6,15H   MEAN URN     )
      WRITE(4,227)SDUR
 227  FORMAT(F10.6,15H   STD DEV URN )
      WRITE(4,229)MNLC
 229  FORMAT(F10.6,15H   MEAN LUN CNT)
      WRITE(4,231)SDLC
 231  FORMAT(F10.2,15H   STD  DEV LC )
      WRITE(4,233)USDEV
 233  FORMAT(F10.4,24H   ASSUM URINALYSIS SDEV)
      WRITE(4,235)MNES
 235  FORMAT(F10.6,15H   MEAN ESIGMA )
      WRITE(4,237)SDES
 237  FORMAT(F10.6,15H   STD DEV ESIG)
      WRITE(4,239)MNEF
 239  FORMAT(F10.6,16H   MN SDEV SIGMA)
      WRITE(4,241)SDEF
 241  FORMAT(F10.6,22H   STD DEV STDEV SIGMA)
C
```

```
      DO 300 K=1,N
      WRITE(2,290)OUTDAT(1,K),OUTDAT(2,K),OUTDAT(3,K),OUTDAT(4,K),
     1 OUTDAT(6,K), OUTDAT(5,K),OUTDAT(7,K),OUTDAT(8,K)
 290  FORMAT(1H ,8F9.5)
 300  CONTINUE
C
 400  CLOSE (1)
      CLOSE (3)
      CLOSE (2)
      CLOSE (4)
C
      STOP
      END
C
C
      SUBROUTINE LRPR(INDAT,OUTDAT,J,I,N,PRIN,PROUT)
C
C*****************************************************************
C* REGRESSION OF INPUT ARRAY INDAT, INDAT(2,I) -LUNG COUNTS- ON  *
C* INDAT(3,I) -SIMULATED URNALYSES-.  OUTPUTS PARAMETERS.        *
C*                                                               *
C*                                                               *
C*                                                               *
C*                                                               *
C*****************************************************************
C
      REAL S,YNFACT,YNEW,XNEW,DF,SSDEVX,INDAT,OUTDAT,ESIG,SIGFAC
      REAL X,Y,Z,SUMX,SUMY,SYSQ,SXSQ,SXY,COUNT,BZERO,BONE,YPRED
      REAL RESD,RESDSQ,SMRESD,SMRESQ,SMSQMN,SMSQAR,SMSQDR,YMN,XMN
      INTEGER I,J,K,N,H
C
      DIMENSION INDAT(3,1000)
      DIMENSION OUTDAT(8,1000)
C
      XMN=0.
      YMN=0.
      RESD=0.
      RESDSQ=0.
      SMRESD=0.
      SMRESQ=0.
      SMSQMN=0.
      SMSQRG=0.
      SMSQAR=0.
      SMSQDR=0.
      BONE=0.
      BZERO=0.
      YI=0.
      COUNT=0.
      SUMX= 0.
      SUMY= 0.
      SYSQ= 0.
      SXSQ= 0.
      SXY= 0.
      XNEW=0.0
      YNEW=0.0
```

```
      YNFACT=0.0
      S=0.0
      DF=0.0
      ESIG=0.
      SIGFAC=0.
C
      COUNT=FLOAT(I)
C
      DO 10 H=1,I
      X=INDAT(3,H)
      Y=INDAT(2,H)
      SUMX=SUMX+X
      SUMY=SUMY+Y
      SXSQ=SXSQ+X*X
      SYSQ=SYSQ+Y*Y
      SXY=SXY+X*Y
 10   CONTINUE
C
      XMN=SUMX/COUNT
      YMN=SUMY/COUNT
      BONE=(SXY-(SUMX*SUMY/COUNT))/(SXSQ-(SUMX*SUMX/COUNT))
      BZERO=(SUMY/COUNT)-BONE*(SUMX/COUNT)
C
      DO 20 K=1,I
      X=INDAT(3,K)
      Y=INDAT(2,K)

      YPRED=BZERO+BONE*X
      RESD=Y-YPRED
      RESDSQ=RESD*RESD
      SMRESD=SMRESD+RESD
      SMRESQ=SMRESQ+RESDSQ
      SMSQMN=SMSQMN+(Y-YMN)*(Y-YMN)
      SMSQAR=SMSQAR+(Y-YPRED)*(Y-YPRED)
      SMSQOR=SMSQOR+(YPRED-YMN)*(YPRED-YMN)


 20   CONTINUE
C
      SSQD=SMRESQ/(COUNT-2.0)
      SSDEVX=(SXSQ-(SUMX*SUMX)/COUNT)
      B0FACT=1.0/SQRT(SSDEVX/((SSQD/COUNT)*SXSQ))
      B1FACT=SQRT(SSQD)/SQRT(SSDEVX)
      S=SQRT(SMRESQ/(COUNT-2.0))
      XNEW=PRIN
      YNEW=BZERO+BONE*XNEW
      YNFACT=S*SQRT(1+(1.0/COUNT)+(((XNEW-XMN)*(XNEW-XMN))/SSDEVX))
      DF=COUNT-2.0
      ESIG=(GAMMA((COUNT-3.0)/2.0)/GAMMA((COUNT-2.0)/2.0))*
     1 SQRT((COUNT-2.0)/2.0)*S
      SIGFAC=SQRT(((COUNT-2.0)*(SMRESQ/(COUNT-2.0))/(COUNT-4.0))
     1 -(ESIG*ESIG))

      OUTDAT(1,J)=BZERO
```

```
      OUTDAT(2,J)=B0FACT
      OUTDAT(3,J)=BONE
      OUTDAT(4,J)=B1FACT
      OUTDAT(5,J)=YNFACT
      OUTDAT(6,J)=YNEW
      OUTDAT(7,J)=ESIG
      OUTDAT(8,J)=SIGFAC
C
      RETURN
      END
C
C
      SUBROUTINE URAND(IY,P)
C
C*********************************************************************
C* GENERATES UNIFORM RANDOM NUMBER BETWEEN 0 AND 1. ALGORITHM        *
C* FROM NETLIB@ORNL.GOV.                                              *
C*                                                                    *
C*********************************************************************
C
      INTEGER IY
      INTEGER IA,IC,ITWO,M2,M,MIC
      DOUBLE PRECISION HALFM
      REAL S,URND
      DOUBLE PRECISION DATAN,DSQRT
      DATA M2/0/,ITWO/2/
      IF (M2 .NE. 0) GO TO 20
      M=1
 10   M2=M
      M=ITWO*M2
      IF (M .GT. M2) GO TO 10
      HALFM=M2
      IA=8*IDINT(HALFM*DATAN(1.D0)/8.D0)+5
      IC=2*IDINT(HALFM*(0.5D0-DSQRT(3.D0)/6.D0))+1
      MIC=(M2-IC)+M2
      S=0.5/HALFM
 20   IY=IY*IA
      IF (IY .GT. MIC) IY=(IY-M2)-M2
      IY=IY+IC
      IF (IY/2 .GT. M2) IY=(IY-M2)-M2
      IF (IY .LT. 0) IY=(IY+M2) + M2
      P=FLOAT(IY)*S
      RETURN
      END
C
      SUBROUTINE  RNDNOR(P,XP)
C*********************************************************************
C* GENERATES A RANDOM NORMAL VARIABLE N(0,1) USING INPUT UNIFORM     *
C* RANDOM VARIABLE FROM URAND, THEN GENERATES A SIMULATED            *
C* URINALYSIS VALUE USING MEAN=INPUT URINALYSIS VALUE -INDAT(1,I)-   *
C* AND STD DEV=USDEV.  USDEV IS INPUT BY USER AT RUNTIME.  A POSTIVE *
C* VALUE OF USDEV IS USED THE STANDARD DEVIATION VALUE.        -EX:  *
C* 1 UNIT OR 10 UNITS-.  A NEGATIVE INPUT VALUE, BETWEEN 0 AND -1,   *
C* INDICATES THAT THE STANDARD DEVIATION SHOULD BE CALCULATED AS THE *
C* GIVEN FRACTION (- SIGN STRIPPED) OF THE INPUT URINALYSIS VAL --.5 *
```

```
C* INDICATES TO CALCULATE STD DEV =1/2  OF EACH MEASURED URINALYSIS  *
C* VALUE.                                                            *
C**********************************************************************
      REAL P,T,C0,C1,C2,D1,D2,D3,XP
      C0=2.515517
      C1=.802853
      C2=.010328
      D1=1.432788
      D2=.189269
      D3=.001308
      COEF=1.
      IF (P .GT. 0.5)  COEF=-1.0
      IF (P .GT. 0.5)  P=1.0-P
      T=SQRT(ALOG(1/(P*P)))
      XP= COEF*(T-((C0+(C1*T)+(C2*T*T))/(1+(D1*T)+(D2*T*T)+(D3*T**3))))
      RETURN
      END
      SUBROUTINE DATSIM(IY,INDAT,I,PRIN,PROUT,USDEV)
C
C
      REAL P,XP,PRIN,PROUT,INDAT
      INTEGER IY,I,K

      DIMENSION INDAT(3,1000)
C
C     P     =RANDOM # FROM UNIFORM DISTRIBUTION
C     URAND IS FUNCTION WHICH RETURNS P, GIVEN SEED IY
C     XP    =NORMAL "Z SCORE" CORRESPONDING TO PROB. OF P
C     K       INDEXES THROUGH INDAT
C
      K=0
      P=0.
      XP=0.
C
      DO 20 K=1,I
 2    CALL URAND(IY,P)
      CALL RNONOR(P,XP)
      INDAT(3,K)=INDAT(1,K)+(XP*USDEV)
      IF (USDEV .GT. 0.) GOTO 5
      INDAT(3,K)=INDAT(1,K)*(1.+(XP*(-1.*USDEV)))
  5   CONTINUE

 20   CONTINUE
      CALL URAND(IY,P)
      CALL RNONOR(P,XP)
      PROUT=PRIN+(XP*USDEV)
      IF  (USDEV .GT. 0.) GOTO 25
      PROUT=PRIN*(1.+(XP*(-1.*USDEV)))
 25   CONTINUE
      RETURN
      END
```

APPENDIX C4

OUTPUT FROM COMPUTER PROGRAMS

FILE: SUMARY 016 A VM/IS

SUMMARY STATISTICS OUTPUT FROM SIMULATION PROGRAM 'SIMU FORTRAN A'.

50.877335	MEAN BO	MEAN OF MAX. LIKELIHOOD ESTIMATES FOR BO
2.828427	STD DEV BO	SD OF MAX. LIKELIHOOD ESTIMATES FOR BO
6.462816	MEAN BOF	MEAN SD OF BO DENSITIES
0.334960	STD DEV BOF	SD OF SD OF BO DENSITIES OVER SIMULATIONS
0.757153	MEAN B1	MEAN OF MAX. LIKELIHOOD ESTIMATES FOR B1
0.265932	STD DEV B1	SD OF MAX. LIKELIHOOD ESTIMATES FOR B1
0.420287	MEAN B1F	MEAN SD OF B1 DENSITIES OVER SIMULATIONS
0.053840	STD DEV B1F	SD OF SD OF B1 DENSITIES OVER SIMULATIONS
46.677032	MEAN YNEWF	MEAN SD OF YN DENSITIES OVER SIMULATIONS
0.410083	STD DEV YNF	SD OF SD OF YN DENSITIES OVER SIMULATIONS
55.420837	MEAN PRED	MEAN OF MODAL VALUES FOR PRED DISTRIBUTIONS
1.259200	STD DEV PRED	SD OF MODAL VALUES FOR PRED DISTRIBUTIONS
102	I, # DATA	# DATA PAIRS IN CALIBRATION DATA SET
1000	N, # SIMU	# SIMULATIONS
10.901917	MEAN URN	MEAN URINALYSIS IN CAL DATA SET
8.937086	STD DEV URN	SD URINALYSIS
59.083740	MEAN LUN CNT	MEAN U-235 LUNG CT IN CAL DATA
47.03	STD DEV LC	SD OF U-235 LUNG COUNTS
-0.5000	ASSUM URINALYSIS SDEV	ASSUMED URINALYSIS SD
46.758789	MEAN ESIGMA	MEAN OF EXPECTATIONS OF SIGMA
0.432651	STD DEV ESIG	SD OF EXPECTATIONS OF SIGMA
3.331172	MN SDEV SIGMA	MEAN OF SD OF DENSITIES OF SIGMA
0.030888	STD DEV STDEV SIGMA	SD OF SD OF DENSITIES OF SIGMA

FILE: SUMARY D15 A VM/IS

```
50.512848   MEAN BO
 2.878977   STD DEV BO
 6.443330   MEAN BOF
 0.348212   STD DEV BOF
 0.780140   MEAN B1
 0.263464   STD DEV B1
 0.412828   MEAN B1F
 0.054744   STD DEV B1F
46.609482   MEAN YNEWF
 0.567424   STD DEV YNF
55.193542   MEAN PRED
 1.369076   STD DEV PRED
    102     I, # DATA
    100     N, # SIMU
10.901917   MEAN URN
 8.937086   STD DEV URN
59.083740   MEAN LUN CNT
    47.03   STD  DEV LC
  -0.5000   ASSUM URINALYSIS SDEV
46.690033   MEAN ESIGMA
 0.568535   STD DEV ESIG
 3.326249   MN SDEV SIGMA
 0.041489   STD DEV STDEV SIGMA
```

FILE: SUMARY 014 A VM/IS

52.455612	MEAN BO
3.417048	STD DEV BO
6.353666	MEAN BOF
0.340539	STD DEV BOF
0.616548	MEAN B1
0.310805	STD DEV B1
0.412595	MEAN B1F
0.055827	STD DEV B1F
46.875671	MEAN YNEWF
0.445853	STD DEV YNF
56.154907	MEAN PRED
1.577016	STD DEV PRED
102	I, # DATA
10	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
-0.5000	ASSUM URINALYSIS SDEV
46.963562	MEAN ESIGMA
0.455007	STD DEV ESIG
3.345727	MN SDEV SIGMA
0.032106	STD DEV STDEV SIGMA

FILE: SUMARY D13 A VM/IS

54.512451	MEAN B0
3.531112	STD DEV B0
6.360235	MEAN B0F
0.309298	STD DEV B0F
0.445397	MEAN B1
0.331735	STD DEV B1
0.430173	MEAN B1F
0.043683	STD DEV B1F
47.130081	MEAN YNEWF
0.467707	STD DEV YNF
57.184814	MEAN PRED
1.545849	STD DEV PRED
102	I, # DATA
5	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
-0.5000	ASSUM URINALYSIS SDEV
47.223663	MEAN ESIGMA
0.471865	STD DEV ESIG
3.364221	MN SDEV SIGMA
0.033829	STD DEV STDEV SIGMA

```
FILE: SUMARY   D12      A   VM/IS

 46.094879   MEAN BO
  0.780134   STD DEV BO
  7.169995   MEAN BOF
  0.163660   STD DEV BOF
  1.192769   MEAN B1
  0.081931   STD DEV B1
  0.508510   MEAN B1F
  0.019977   STD DEV B1F
 46.299438   MEAN YNEWF
  0.001000   STD DEV YNF
 53.252014   MEAN PRED
  0.100050   STD DEV PRED
     102     I, # DATA
    1000     N, # SIMU
 10.901917   MEAN URN
  8.937086   STD DEV URN
 59.083740   MEAN LUN CNT
     47.03   STD  DEV LC
   -0.1000   ASSUM URINALYSIS SDEV
 46.357254   MEAN ESIGMA
  0.001000   STD DEV ESIG
  3.302564   MN SDEV SIGMA
  0.001000   STD DEV STDEV SIGMA
```

FILE: SUMARY D11 A VM/IS

46.020966	MEAN B0
0.897176	STD DEV B0
7.158410	MEAN B0F
0.137620	STD DEV B0F
1.196708	MEAN B1
0.084709	STD DEV B1
0.505969	MEAN B1F
0.019704	STD DEV B1F
46.280731	MEAN YNEWF
0.189697	STD DEV YNF
53.201050	MEAN PRED
0.437365	STD DEV PRED
102	I, # DATA
100	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
-0.1000	ASSUM URINALYSIS SDEV
46.338425	MEAN ESIGMA
0.192996	STD DEV ESIG
3.301215	MN SDEV SIGMA
0.015625	STD DEV STDEV SIGMA

FILE: SUMARY D10 A VM/IS

46.582199	MEAN BO
1.066984	STD DEV BO
7.150089	MEAN BOF
0.138389	STD DEV BOF
1.152836	MEAN B1
0.096166	STD DEV B1
0.507741	MEAN B1F
0.020729	STD DEV B1F
46.374039	MEAN YNEWF
0.190941	STD DEV YNF
53.499191	MEAN PRED
0.511585	STD DEV PRED
102	I, # DATA
10	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
-0.1000	ASSUM URINALYSIS SDEV
46.433685	MEAN ESIGMA
0.190941	STD DEV ESIG
3.307963	MN SDEV SIGMA
0.013656	STD DEV STDEV SIGMA

FILE: SUMARY D9 A VM/IS

47.073425	MEAN B0
1.002926	STD DEV B0
7.183200	MEAN B0F
0.111858	STD DEV B0F
1.117676	MEAN B1
0.086685	STD DEV B1
0.515670	MEAN B1F
0.013165	STD DEV B1F
46.472916	MEAN YNEWF
0.190086	STD DEV YNF
53.779434	MEAN PRED
0.492126	STD DEV PRED
102	I, # DATA
5	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
-0.1000	ASSUM URINALYSIS SDEV
46.533615	MEAN ESIGMA
0.190086	STD DEV ESIG
3.315039	MN SDEV SIGMA
0.014084	STD DEV STDEV SIGMA

```
FILE: SUMARY    D4         A    VM/IS

 45.986618    MEAN B0
  0.531302    STD DEV B0
  7.180245    MEAN B0F
  0.123758    STD DEV B0F
  1.201644    MEAN B1
  0.057657    STD DEV B1
  0.509395    MEAN B1F
  0.009522    STD DEV B1F
 46.289139    MEAN YNEWF
  0.001000    STD DEV YNF
 53.197052    MEAN PRED
  0.001000    STD DEV PRED
    102       I, # DATA
    1000      N, # SIMU
 10.901917    MEAN URN
  8.937086    STD DEV URN
 59.083740    MEAN LUN CNT
     47.03    STD  DEV LC
    1.0000    ASSUM URINALYSIS SDEV
 46.346466    MEAN ESIGMA
  0.001000    STD DEV ESIG
  3.301785    MN SDEV SIGMA
  0.001000    STD DEV STDEV SIGMA
```

FILE: SUMARY D3 A VM/IS

45.979324	MEAN B0
0.676070	STD DEV B0
7.187636	MEAN B0F
0.070789	STD DEV B0F
1.200077	MEAN B1
0.059168	STD DEV B1
0.509215	MEAN B1F
0.006522	STD DEV B1F
46.293076	MEAN YNEWF
0.172255	STD DEV YNF
53.179688	MEAN PRED
0.371830	STD DEV PRED
102	I, # DATA
100	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
1.0000	ASSUM URINALYSIS SDEV
46.349915	MEAN ESIGMA
0.152835	STD DEV ESIG
3.302021	MN SDEV SIGMA
0.014046	STD DEV STDEV SIGMA

FILE: SUMARY D2 A VM/IS

```
46.491028    MEAN BO
 0.606318    STD DEV BO
 7.165442    MEAN BOF
 0.056816    STD DEV BOF
 1.161008    MEAN B1
 0.048280    STD DEV B1
 0.509475    MEAN B1F
 0.006354    STD DEV B1F
46.371887    MEAN YNEWF
 0.110240    STD DEV YNF
53.457047    MEAN PRED
 0.332682    STD DEV PRED
     102     I, # DATA
      10     N, # SIMU
10.901917    MEAN URN
 8.937086    STD DEV URN
59.083740    MEAN LUN CNT
    47.03    STD  DEV LC
   1.0000    ASSUM URINALYSIS SDEV
46.430969    MEAN ESIGMA
 0.112191    STD DEV ESIG
 3.307817    MN SDEV SIGMA
 0.008337    STD DEV STDEV SIGMA
```

FILE: SUMARY DI A VM/IS

46.813049	MEAN BO
0.515388	STD DEV BO
7.175302	MEAN BOF
0.063948	STD DEV BOF
1.137032	MEAN B1
0.038824	STD DEV B1
0.512737	MEAN B1F
0.004367	STD DEV B1F
46.432693	MEAN YNEWF
0.062500	STD DEV YNF
53.635193	MEAN PRED
0.296464	STD DEV PRED
102	I, # DATA
5	N, # SIMU
10.901917	MEAN URN
8.937086	STD DEV URN
59.083740	MEAN LUN CNT
47.03	STD DEV LC
1.0000	ASSUM URINALYSIS SDEV
46.492661	MEAN ESIGMA
0.062500	STD DEV ESIG
3.312164	MN SDEV SIGMA
0.005167	STD DEV STDEV SIGMA

FILE: FOROUT DATA A VM/IS

```
XMN,YMN,BO,B1
     10.901960       59.084137       45.795166        1.218951
SUMX,SUMY,SYSQ,SXSQ,SXY
   1112.000000    6026.582030  579466.187000   20190.000000   75534.812500
SMRESQ,COUNT,SSDEVX,S
 211401.750000     102.000000    8067.023437       45.978444
BOFACT,B1FACT
      7.202208        0.511915
ESIG,SIGFAC
     46.326913        3.313262
SIGMOD
     45.750262
```

FILE: PRDIN DATA A VM/IS

6.00

FILE: PROJUT DATA A VM/IS

6.000000	53.108872	100.000000	46.271375

FILE: PARMS DATA A VM/IS

48.00192	6.71419	1.00366	0.46939	54.02388	46.50545	46.57042	3.31766
49.73906	6.25418	0.81794	0.37356	54.64670	46.44014	46.52098	3.31423
49.16710	6.34590	0.99423	0.44186	55.13246	46.36981	46.46146	3.30992
47.58839	6.73415	1.01818	0.43899	53.69749	46.32513	46.39160	3.30512
51.83929	6.23300	0.67399	0.39011	55.88319	46.83839	46.92789	3.34320
49.64496	6.80893	0.94751	0.50399	55.33003	46.71997	46.80313	3.33417
54.36703	7.02265	0.39166	0.48754	57.21700	47.39958	47.47015	3.38190
53.75015	6.92862	0.48713	0.46870	56.67293	47.29750	47.36794	3.37449
48.51984	6.79236	0.96828	0.45979	54.32951	46.53044	46.60092	3.31983
50.17500	6.21841	0.76030	0.35911	54.73682	46.51031	46.59026	3.31920
54.40941	6.63417	0.41139	0.41580	56.87776	47.31685	47.39168	3.37612
51.34230	6.25689	0.69367	0.37995	55.51028	46.76331	46.84877	3.33765
49.28014	6.35338	0.92279	0.41558	54.81088	46.40549	46.49068	3.31217
51.99803	5.88590	0.66893	0.34720	56.01167	46.65457	46.76305	3.33139
48.27779	5.79725	0.88218	0.29946	53.57083	45.59758	45.68199	3.25443
50.41003	6.19599	0.75993	0.36542	54.96964	46.54222	46.62553	3.32166
49.99998	6.20920	0.85902	0.39662	55.14810	46.45683	46.54636	3.31600
50.71104	6.60141	0.80049	0.45191	55.51398	46.81041	46.89310	3.34094
50.14253	5.89340	0.75882	0.31769	54.69546	46.23201	46.32001	3.29980
50.61714	6.73998	0.80636	0.46838	55.45828	46.85291	46.93179	3.34342
53.46245	7.05726	0.52093	0.49163	56.58804	47.28938	47.35800	3.37377
49.36528	6.42036	0.94292	0.43707	55.02281	46.46503	46.55205	3.31622
49.44281	6.93637	0.88677	0.47742	54.76343	46.75429	46.82227	3.33560
52.70872	6.32841	0.61238	0.41443	56.38300	47.02040	47.11154	3.35639
51.96347	6.90621	0.65395	0.47018	55.88715	47.09836	47.16904	3.36021
47.25661	7.09911	1.08153	0.49656	53.74577	46.47171	46.53214	3.31501
53.88112	5.86565	0.44012	0.30860	56.55785	47.03885	47.13312	3.35767
54.40698	6.37412	0.49061	0.46255	57.38563	47.25012	47.35100	3.37348
49.15364	6.35514	0.84980	0.37827	54.25241	46.38991	46.46493	3.31003
46.87468	6.51318	1.09453	0.41967	53.44186	46.00856	46.08157	3.28281
48.62682	6.01974	0.84459	0.32061	53.69437	45.97354	46.05182	3.28095
46.57690	6.64368	0.93310	0.42803	54.17552	46.45979	46.53030	3.31493
50.74110	5.76651	0.68328	0.28963	54.84076	46.26100	46.35063	3.30213
52.50603	6.47583	0.59411	0.40881	56.07068	47.04698	47.12798	3.35752
53.25266	6.07004	0.54317	0.36910	56.54167	47.01086	47.10637	3.35585
54.04294	6.79274	0.46748	0.45868	56.84779	47.30122	47.37769	3.37507
49.71646	6.46995	0.96574	0.47067	55.51086	46.55838	46.65125	3.32343
51.43500	6.50845	0.78353	0.46998	56.13614	46.88195	46.97478	3.34652
51.45335	6.39390	0.75820	0.44003	56.00252	46.83945	46.93149	3.34349
52.66335	6.85837	0.54840	0.43118	55.95374	47.17973	47.24251	3.36547
53.65138	6.53416	0.50347	0.42989	56.70221	47.20921	47.29341	3.36932
44.59943	6.30205	1.23491	0.38036	52.00391	45.22629	45.29553	3.22674
50.72845	6.31756	0.74456	0.44705	55.19583	46.90784	46.97600	3.34678
53.40239	6.81093	0.57506	0.50372	56.85275	47.22905	47.31580	3.37077
46.75810	6.69082	1.18403	0.47218	53.86229	46.11577	46.19293	3.29087

56.91187	6.87390	0.20996	0.48770	58.17163	47.49841	47.57903	3.38954
57.98679	6.09525	0.11016	0.39226	58.64774	47.50177	47.60432	3.39124
50.42345	6.28294	0.81656	0.40477	55.32280	46.59445	46.68265	3.32578
51.00125	6.16790	0.78465	0.39972	55.70912	46.63731	46.73126	3.32915
56.23901	6.37620	0.30565	0.46641	58.07294	47.41846	47.52110	3.38540
49.16467	6.87410	0.93138	0.48012	54.75294	46.67921	46.75151	3.33076
51.90778	5.98637	0.68530	0.36586	56.01959	46.71179	46.80901	3.33472
48.97581	6.68527	0.91196	0.43919	54.44759	46.55656	46.62836	3.32185
52.75793	6.01961	0.50998	0.31126	55.81783	46.91238	46.99637	3.34805
49.18835	6.17015	0.89987	0.37867	54.58757	46.24686	46.33272	3.30080
47.58025	6.63702	1.06615	0.44760	53.97714	46.25212	46.32697	3.30050
50.66096	6.01126	0.75199	0.34818	55.17290	46.45996	46.54979	3.31640
48.85794	6.11233	0.81049	0.32434	53.72084	46.12807	46.20233	3.29135
46.20844	6.92740	1.09352	0.44408	52.75955	46.18860	46.24168	3.29436
54.92233	6.24628	0.37091	0.37094	57.14720	47.29912	47.38684	3.37594
47.12186	6.40641	1.09219	0.41406	53.67499	45.96996	46.04800	3.28073
49.16908	7.06375	0.95215	0.51465	54.88197	46.75771	46.82640	3.33611
55.59766	7.20949	0.34103	0.53722	57.64386	47.45387	47.52742	3.38594
51.72583	6.35585	0.60026	0.35652	55.32741	46.83889	46.96205	3.34565
50.26120	6.80836	0.76680	0.43557	54.86198	46.83717	46.90194	3.34152
56.85548	6.74971	0.20147	0.44013	58.06430	47.49760	47.57324	3.38926
52.47531	7.01665	0.57331	0.45635	55.91518	47.19208	47.25165	3.36634
43.04713	5.81489	1.44862	0.35101	51.73886	43.93910	44.02139	3.13500
53.78152	5.74866	0.50475	0.32507	56.81000	46.95541	47.05910	3.35268
50.66708	6.40358	0.78403	0.41510	55.37126	46.71123	46.79568	3.33381
51.60561	5.92123	0.66747	0.33397	55.61043	46.60640	46.69952	3.32699
50.34383	6.31564	0.80751	0.40102	55.18889	46.60057	46.68594	3.32592
49.91016	6.26692	0.81721	0.38143	54.81343	46.48386	46.56633	3.31747
49.28568	6.44495	0.81785	0.37894	54.19275	46.48135	46.55122	3.31637
49.88197	6.49786	0.80581	0.40285	54.71684	46.62360	46.69798	3.32680
51.22289	6.53815	0.74752	0.44050	55.70798	46.86781	46.95180	3.34488
53.12555	6.53900	0.53246	0.41162	56.32028	47.15099	47.22952	3.36479
46.85785	6.31266	1.06226	0.38374	53.23140	45.82245	45.89703	3.26973
51.62447	6.29230	0.70565	0.40508	55.85838	46.82701	46.91653	3.34240
48.96181	6.64578	0.91335	0.43454	54.44188	46.53185	46.60472	3.32016
50.30220	6.16147	0.84052	0.39477	55.34534	46.48694	46.57899	3.31846
47.33696	6.37084	1.10025	0.41985	53.93846	45.98553	46.06746	3.28196
49.08684	5.30892	0.83311	0.30467	54.08549	45.84949	45.93652	3.27276
47.38593	6.23346	1.02958	0.37821	53.56340	45.87213	45.95090	3.27373
52.03168	6.73403	0.63008	0.43671	55.81216	47.06297	47.13503	3.35778
52.85703	6.03954	0.57158	0.35695	56.28650	46.93005	47.02397	3.35010
50.66817	6.40674	0.78833	0.41784	55.39816	46.71251	46.79747	3.33388
46.45662	6.56982	1.14338	0.43140	53.31693	45.96140	46.03363	3.27958
47.18268	6.58177	1.10970	0.44429	53.84085	46.12721	46.20364	3.29172
47.69151	6.50946	0.98016	0.40085	53.57243	46.19087	46.26022	3.29580
47.21732	6.33621	1.08498	0.40589	53.72722	45.92772	46.00768	3.27753
49.36542	6.82846	0.83363	0.43323	54.36722	46.70375	46.76514	3.33157
51.44862	6.60280	0.68979	0.42619	55.58737	46.93460	47.01128	3.34904
48.86160	6.19499	0.85895	0.35349	54.01529	46.19911	46.27641	3.29695
49.45070	6.17164	0.88240	0.38102	54.74507	46.30836	46.39517	3.30538
49.17291	6.14154	0.90284	0.37573	54.58994	46.21837	46.30501	3.29898
55.53162	6.45908	0.29660	0.37298	57.31122	47.39803	47.47321	3.38197
41.98897	6.92252	1.75156	0.54298	52.49834	45.24756	45.32332	3.22897
52.18712	6.05794	0.61130	0.34972	55.85490	46.82207	46.91176	3.34218
47.64093	6.54650	1.09041	0.44874	54.18338	46.19604	46.27641	3.29684

APPENDIX D

CORRECTIONS TO CALIBRATION DATA SET

APPENDIX D

CORRECTIONS TO CALIBRATION DATA SET

A difficulty often encountered in the measurement of radioactivity at near-background levels is that the sample net counting rate may be negative due to random fluctuations in the source and background emissions. Little (1982) points out that, when it is known that the net counting rate must be positive, there are approaches for estimating the net counting rate which incorporate this prior information.

These approaches result in probability density functions for the net count rate which contain zero probability for negative net count rates. The expectation and mode of these densities can be used as estimates of the true net counting rates; these quantities are always positive.

Eleven of the record pairs in the calibration data set utilized in this thesis initially contained negative estimates for Uranium-235 lung burden. Smith et. al. (1987) indicate that the Uranium-235 lung burden estimates were calculated using gross counts of gamma emissions in the 186 keV range along with background count data. They further indicate that this process could and did sometimes result in negative Uranium-235 lung burden estimates.

It would be appropriate to apply the techniques suggested by Little to correct the negative Uranium-235 lung burden estimates found in the calibration data set. However, these techniques require information concerning the source and background counting rates and times which was not available. Therefore, it was simply recognized that negative Uranium-235 lung burdens were physically impossible and, as a first approximation toward correcting them, the negative values were set to zero.

VITA

Bradley Wicker Dolan was born in San Francisco, California on December 2, 1956. He attended elementary and high school in Friendsville, Tennessee and was graduated from Friendsville High School in May 1974. In September of 1974, he entered the University of Tennessee but left before attaining an undergraduate degree. During the following years, he worked in nuclear power plant operations as an employee of the Tennessee Valley Authority at Browns Ferry Nuclear Plant and as a consultant at other power plants. During this time he completed substantial training and coursework in topics relating to radiological protection and nuclear engineering. In June 1985, he returned full-time to the University and in December 1986 he received a Bachelor of Science degree in Operations Management.

After accepting a position in 1987 with the hazards assessment group at the Oak Ridge Associated Universities' Center for Epidemiologic Research, he enrolled in the Radiation Protection Engineering Master's program at the University of Tennessee. He will receive the Master's degree in Nuclear Engineering in May 1992.

The author is a member of the Health Physics Society and the Reactor Operator Society. Mr. Dolan is currently employed by Science Applications International Corporation.