Order Number 9034353

Computer-aided localization of neurological diseases

Parlar, Yusuf, D.Eng.Sc.

New Jersey Institute of Technology, 1990

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Computer Aided Localization of Neurological Diseases

by

Yusuf Parlar

Dissertation submitted to the Faculty of the Graduate School
of the New Jersey Institute of Technology in partial fulfillment
of the requirements for the degree of
Doctor of Engineering Science
1990

APPROVAL SHEET

Title of Thesis: Computer Aided Localization of Neurological Diseases

Name of Candidate: Yusuf Parlar
Doctor of Engineering Science, 1990

Thesis and Abstract Approved: _____ Date _____
Dr. Andrew U. Meyer
Professor
Department of Electrical and Computer Engineering

Signature of other members of the thesis committee. _____ Date _____
Dr. Rose A. Dios
Associate Professor
Deparment of Mathematics

_____ Date _____
Dr. Stanley Reisman
Professor
Department of Electrical and Computer Engineering

_____ Date _____
Dr. Peter Engler
Professor
Department of Electrical and Computer Engineering

_____ Date _____
Dr. William K. Weissman
Adjunct Research Professor
Department of Electrical & Comp. Eng.

VITA

Name: Yusuf Parlar

Degree and date to be conferred: D. Eng. Sc., 1990.

Secondary education: Bahçelievler Deneme Lisesi,
Turkey, 1974

Collegiate institutions attended:	Date	Degree	Date of Degree
New Jersey Institute of Technology	9/83-5/90	D. Eng. Sc.	May 1990
Polytechnic Institute of New York	9/82-5/83	M.S.E.E	May 1983
Middle East Technical University	9/74-2/81	B.S.E.E	May 1981

Major: Electrical Engineering.

Positions held: Teaching Fellow, New Jersey Inst. of Tech.,
Newark, NJ, 9/87-5/90
Teaching Assistant, New Jersey Inst. of Tech.,
Newark, NJ, 9/83-5/87
Teaching Assistant, Polytechnic Institue of New York,
Brooklyn, NY, 1/83-5/83
Research Engineer, Clarke-Hess Comm. Res. Corp.
New York, 1/83-5/83
Teaching Assistant, Middle East Technical University,
Ankara, Turkey, 9/81-5/82

ABSTRACT

Title of Thesis:	Computer Aided Localization of Neurological Diseases
Yusuf Parlar	Doctor of Engineering Science. 1990
Thesis directed by:	Prof. Dr. Andrew U. Meyer
	Assoc. Prof. Dr. Rose A. Dios

Computers in medicine has brought new dimensions and better understanding of uncertainties in the field of medical sciences in the last decade. This thesis is focused upon the localization of neurological lesions in the human nervous system. It relates clinical neurological test outcomes to pathways of function or malfunction. Certain methods are proposed–empirical. stochastic. deterministic–to estimate the spatial distributions of lesion probabilities.

First. a Bayesian model is presented to estimate the posterior probability of lesion from á priori information, based on the test outcomes. Due to unavailable data alternative methods and models are presented: Regression Analysis. Monte Carlo simulation, and finally a new model known as Logistic Sigmoid Nonlinearity is proposed for probability estimation. This dissertation analyzes each of these models and alternative methodologies in detail.

Sevgili büyüklerim için

To my Parents

Acknowledgement

I would like to express my appreciation to Dr. Andrew U. Meyer for his support, knowledge and insight. I would also like to express my thankfulness to Dr. William Weissman for contributing his plethora of neurological knowledge. I would also like to profusely thank Dr. Rose A. Dios for her constant help, valuable guidance and support. without whom this dissertation would not have been completed.

Many thanks are due to the distinguished members of the committee: Dr. Stanley Reisman and Dr. Peter Engler. Their encouragement and valuable discussions have improved the quality of this dissertation.

The author would also like to thank the valuable members of the Electrical and Computer Engineering department for their help and continuous support.

I also would like to extend a special thanks to my immediate family. Many thanks go to my mother and my brother and sisters and to my wife Mary, and our son Erin. for their selfless dedication and generous support, far in excess of reasonable expectations.

Contents

List of Tables

List of Figures

Chapter 1

INTRODUCTION

1.1 Quantifying the Clinical Neurological Examination

The human nervous system is an extremely involved structure responsible for highly complicated functions and activities such as control and regulation of body processes (e.g., temperature, CO_2 concentration,ph level), consciousness, abstract thought, memory, and the interpretation of emotion, etc. The basic unit of the nervous system is the individual nerve cell, or neuron. Only about 10 percent of the cells in the nervous system are neurons, the remainder are glial cells, which sustain the neurons metabolically, support them physically, and help regulate the ionic concentrations in the extracellular space. Neurons occur in many different shapes and sizes, but they can be considered as consisting of three basic parts: (1) the dentrites and cell body, (2) the axon, and (3) the axon terminals. Regardless of their shape, neurons can be divided into three functional classes: afferent neurons, efferent neurons, and interneurons. *Afferent neurons* carry information from receptors into the brain or spinal cord. *Efferent neurons* transmit the final integrated information from the central nervous system out to the effector organs(muscle or glands). *The interneurons*, which both originate and terminate within the central nervous system, account for the 99 percent of all nerve cells [10].

This study is concerned with the localization of lesions in the human nervous

system. In medicine, the term *lesion* is used to describe a broad number of conditions involving "an alteration of structure or of functional capacity due to injury or disease, or any structural perversion, which produces or maintains discomfort or functional disorder, or impairs natural immunity of the body or a part" ([5]). In this work the term lesion is used to imply functional disorder or discomfort.

In most cases, an experienced neurologist or a neurosurgeon might be able to locate lesions in the nervous system using only "routine" clinical neurological examinations. These clinical neurological examinations do not require expensive equipment, such as medical imaging instrumentation including Radiography , Computer Aided Tomography (CAT), Positron Emission Tomography (PET), etc., which also subject the patient to hazardous radiation. However interpretation of these clinical examinations require a thorough knowledge of neuroanatomy and neurophysiology, which a nonspecialist medical practitioner need not have. Clearly this is an ideal situation in which a computer equipped with an intelligent program and an adequate database might prove to be useful.

The application of computers to medicine has increased sharply during the last decade. This may be attributed to the recent spectacular advances in technology which made computers faster, more reliable and–more importantly–cheaper. The field of neurology has also benefited from these developments as evidenced by a sizable number of publications on the application of computers in this field. One of the pioneering works in this area is the program developed by Meyer and Weissman [19], [20] used to locate lesions in the brainstem [1973].

In the aforementioned program the brainstem was divided into 10 sections with each section, in turn, subdivided into 100 volume units. Each volume unit was associated with various neural–pathways. A signal flow analysis through every known neural–pathway due to any test outcome was then carried out. Each volume unit was coded and used to find its involvement in pathways of malfunction (or

function), which is called the malfunction factor (or function factor), depending on the test outcome observed. In addition to these factors the modified malfunction and function factors were computed. These factors were then displayed in their proper locations to identify possible lesions.

In this study, it is proposed to extend the work of Meyer and Weissman to develop a variety of approaches for localization of lesions in the human nervous system using the outcomes of clinical neurological tests. For the sake of brevity, the system to be developed will be referred to as **CALOND.** [1] It will be concerned with the examination of function or malfunction of neural–pathways which, indeed, represents the essence of clinical neurology. The system will relate clinical test outcomes to pathways of probable malfunction, as well as probable function. Then the test outcomes will be used to find the conditional probability of a malfunctioning **voxel** for a given set of test outcomes and then the findings will be shown on a map of sections of the nervous system to indicate the regions of malfunction as well as function. The purpose of this study is to propose and enhance a variety of approaches to be used in localization of lesions in the human nervous system through computer analysis. The general strategy will focus upon utilizing the relation between test outcomes (of clinical neurological tests) and neural pathways to locate sites of probable lesion. In addition to its clinical use, CALOND will be designed for use as a teaching tool in the neurosciences.

The following section will cover the current research on Computer–Aided Medical Diagnosis with the emphasis in the field of neurology. Some related studies emphasizing other domains of medicine, will also be included, illustrating the breadth of computer applications to this field.

[1] The name CALOND represents the initials of "Computer Aided Localization Of Neurological Diseases".

1.2 Computer-Aided Medical Diagnosis

This section presents an overview of the literature on Computer–Aided Medical Diagnosis, with attention focused mainly on papers involving applications related to neurological anatomy. A brief discussion of peripherally related literature is also included because it is considered by the author as worthwhile in analyzing programming and/or system structures.

Du Boulay [11] developed a system that used information of neuroradiologic tests to determine intracranial tumors through the use of a simple weighting technique. Clinically suspected tumors were divided into three groups and diagnostic methods were applied separately for each group. Two main programs were written: one added information about previous patients to the main file; the other program suggested diagnoses in decreasing order of probabilities as well as recommended subsequent, confirmatory testing for new patients. His conclusion was that his approach was insufficient for the analysis undertaken and his sentiment is that the computer will in no way replace the physician [1968].

Wortman [35] developed an information processing system that was used to simulate the diagnostic behavior of the physician. Information about diseases and related symptoms of the cerebellar syndrome were selected by a neurologist and the program was tested against the clinician during an interactive session. There was a consistency between the neurologist's final approach and the system's diagnosis. For this system the disease area was limited and all testing was simulated.

Mori [21] and his associates developed a system that was used for the differential diagnosis of brain lesions with the use of information assembled from neuroradiologic tests. 240 true positive brain scans were used and 86 scan parameters (density, shape, number, location, etc.) were extracted from these scans without any reference to neurological signs and symptoms and then the maximum likelihood method was applied with 77% accuracy [1975].

Wiener's [34] system was based upon the logical relations between a disease and its associated clinical findings . The system was used for diagnosing the comatose patient. For each disease, related clinical findings were separated with respect to representation of degree of diagnostic certainty, and consistency with the sequence in which the findings became known. Then threshold logic was applied, expressing boolean combinations of findings sufficiently to confirm a given diagnostic stage [1975].

Stewart [30] and Cala developed a mathematical method for diagnosis of site and type of intracerebral mass lesions. Data for a new patient was coded and entered, and the number of basic test results that were common to both the new patient and the 'past–patient' were calculated, and various weightings were given for positive results that were common to new and 'past–patients'. Basically it was an application of Bayesian statistics and conditional probabilities in which identifying the diagnosis with the largest probability of occurrence was conditional upon the observations (based upon those tests applied) [1975].

Okada [23] and his associates developed a system that was using a maximum likelihood method for the differential diagnosis of multiple sclerosis. Their program consisted of five parts; entering new patient data; renewing or correcting the previously recorded information; retrieval of information; computation of parameters required for automated diagnosis; and diagnosing patient on the basis of computed and stored parameters. The system was designed for a limited domain and thus expansion to other domains was not possible [1977].

In his next approach Du Boulay [12] and his associates again divided the cerebral tumors into three groups and applied different diagnosis methods to each group separately. Rather than of using weighted scoring, by defining $D_i(i = 1,..k)$ as diseases and $S_j(j = 1,..n)$ as symptoms, they used Bayes' theorem to find the conditional probabilities $P(D_i/S)$ to use in diagnosis. Assuming that symptoms

are independent within each disease one can write

$$P(D_i/S) = P(S/D_i)P(D_i)/P(S);$$

and if

$$S = \bigcap_{k=1}^{n} S_k$$

then

$$P(S/D_i) = P(S_1/D_i)P(S_2/D_i)\ldots P(S_n/D_i)$$

Thus the conditional probability $P(S_j/D_i)$ has to be estimated using prior patient data. For this the following observed frequency was calculated:

$$f_{ji} = M_{ji}/(M_{ji} + N_{ji})$$

where M_{ji} is the number of cases of disease i in which jth sign was present and N_{ji} is the number of cases of disease i in which jth sign was absent. The accuracy was compared with the original study using the weighted scoring technique and also with just the radiologist's diagnosis. For the first group of patients, the computer results confirmed those of radiologists, in fact, this was possibly used for teaching purposes. For the second group of patients, the results were more accurate than that of the radiologist's diagnosis. Third group patients were tested in relation to vertebral angiography and it was concluded that the support of some clinical evidence or radiological tests was necessary [1977].

A decision guide for meningitis in children was developed by Knapp and his associates [17]. 193 cases were reviewed and statistically analyzed to determine optimal clinical discriminators for the disease. Numerical weights were then assigned to various signs and symptoms by a statistical technique with the constraint that the sum of the weights for all symptoms present would generate the the discriminant equation for the diagnosis of meningitis [1977].

J. A. Reggia [25] developed a production rule system to localize central nervous system lesions in unconscious patients. In general, production rule systems

constituted a programming methodology for modeling symbol–processing aspects of recognition. The system has a database that includes a set of rules and a rule interpreter and selector. The rule interpreter may be antecedent–driven, where the occurrence of one or more of antecedents triggers the application of the rule inferring its consequences; or the rule interpreter may be consequent-driven, where the interpreter selects a rule with a fact to be established as a consequent and then tries to verify it. For neurological localization, first an examination is conducted generating the data. then the data are analyzed to determine the site(s) of brain damage most likely to explain the examination findings. In the case when the patient is in a coma of unknown etiology, it is a critical situation since different disease processes that cause coma may involve different regions of the nervous system. Reggia's database included dynamic knowledge about the patient and consisted of a set of attributes possessed by the patient. These are attributes examination based and inferred. The system applied the MYCIN like rule-based program. MYCIN will be explained in detail in the current section. It follows the IF–THEN format to express rules. The control structure used is consequent driven and produces a search and/or goal tree.The program begins with the start goal. and then sets up subgoals. These in return may set up more subgoals or result. The program was tested on simulated patients for four different categories of unconscious patient. As a result Reggia gave the following observations:

- Expressing neurological localization knowledge as a collection of rules is very difficult.
- A collection of rules is not a good model of the organization of neurological localization knowledge as used by the physician.
- The interpretation of neurological examination abnormalities is highly context–dependent, and this may lead to combinatorial problems.

- One way to improve the compactness and understandability of rules is to create and use domain-specific 'macropredicates'[1978].

Catanzarite [7] had developed a computer program for localization and diagnosis in clinical neurology which was called "NEUROLOGIST". The program consisted of four modules. The 'Input' module starts with user orientation, then history and physical examination data are entered for which findings are mapped onto 'status' for each of the 100 nervous system tracts represented in the system. The 'Loc' module is then used to localize the lesions. The lesions may be anatomical, biochemical, or physiologic. For the anatomical lesions a drawing based localization is used and for the biochemical and physiologic lesions a rule based localization is used. After the localization, the program checks whether the findings are indeed consistent with a lesion at this localization, and shows percentages of findings as explained by lesions at this locus. The 'Hgen' module selects lesions which best explain observed 'malfunction' and uses the location of the lesion, together with the mode of disease onset, to retrieve a list of tentative diagnostic hypotheses from the hypothesis generating table. The 'Hypothesis testing' modules database consists of disease specific information. The final stage gives an evaluation of diagnostic hypotheses, and provides explanations. The network feature of hierarchical structure is used for disease representation. In the same module the program rescores all diagnoses, and the highest scoring diagnosis – not investigated as of yet – becomes the active hypothesis. When all diagnostic hypotheses have been investigated, a diagnostic summary is given [1980].

LOCALIZE is a computer program developed by M. B. First [13] and his associates to assist physicians with localization of lesions in the peripheral nervous system. The input to the system is clinical and consists of electromyographic evidence of specific muscle weaknesses. The program's database was constructed from neuroanatomic references and shows the interconnections of peripheral nervous sys-

tem components. The database was represented as a network. It has 2224 named nervous systems structures and 9796 links among them. The program starts with data collection, in particular with the identification of clinically weak muscles. Then the program responds with the review of the most commonly tested muscles to check that nothing has been omitted. After the data entry, the program enters the localization of lesions phase. First the nerve segments that participate in the supply of the affected muscles are identified, then fibers which supply each affected muscle, proximally to the spinal cord, are traced and updated in the knowledge base. Any set which includes at least one highlighted segment from each traced pathway will account for all of the deficits. Solution sets that consist of lesions which anatomically lie most distal and proximal are constructed. First taking into account most distal lesions, the program generates alternative solution sets by replacing set elements with more proximal lesion sites from the highlighted pathways. Then by applying a convergence algorithm the number of hypothesized lesion sites are reduced, excluding multiple lesions. Once the proper convergence point is found, to apply the substitution, consistency checks must be satisfied. If the consistency check fails three times, re–examination of the muscle will be carried out. For plexus lesions a different approach is used. In addition to the procedure described above the 'plexus algorithm' is used to determine the validity of a plexus solution. Each peripheral nerve in the solution set is followed until either it diverges or the plexus is reached. In the final stage, the most proximal site for the occurrence of the lesions, consistent with the findings, are then determined. Sensory deficits and reflex changes are not included in the system despite the fact that they could increase the accuracy of the program [1982].

NEUREX (Neurologic Expert) is a diagnostic expert system developed by Xiang and his associates [36].In this system knowledge of the spatial structure and function of neuroanatomy is represented as a semantic network, in which every

cross section and every region represents an anatomic concept. Connectivity of segments is asserted between corresponding concepts. Each tract is represented by an atomic node. Anatomically significant components of the Central Nervous System, Peripheral Nervous System, and transverse nerve segments of the Peripheral Nervous System are represented by unique atomic nodes and connectivity relations are specified by nodes with proximal and distal arcs [1986].

The remainder of this discussion will focus upon works related to Computer-Aided Medical Diagnosis in other fields of medicine.

One of the earliest work in medical diagnosis using computers was done in the late 1950's by Ledley and Lusted [18]. They have illustrated the automated decision use of computers in the diagnosis of congenital heart disease and discussed the potential value of probabilistic models for diagnostic inference. [1959]

Warner and his associates [32] developed a computer program used for the diagnosis of congenital heart disease based upon Bayes' Theorem. It is assumed that symptoms are independent of each other within a given disease, and that diseases are mutually exclusive. Data compiled from patients were used to generate a symptom-disease matrix consisting of 53 symptoms and 35 disease entities. Then, based upon the presence or absence of these symptoms in a new patient, the program's diagnosis was compared to that of two experienced physicians. It was found that the system's accuracy was equal to that of expert in that field. Also the accuracy improved with refinements in the data matrix [1964].

Gorry and Barnett [14] suggested that Warner's program would not be feasible for many applications since it required determination of 53 observations for every patient to be diagnosed: therefore sequential diagnosis was proposed through the use of a modified Bayes' Theorem. They defined an *attribute* to be a sign or symptom which can provide information for the diagnosis: a *test* as the means employed to detect the presence or absence of one or more attributes: and the selection of a test or sequence of tests as the *test selection function.* The program's information base constituted the medical "experience" of the program. The inference function was used to construct the current view of the diagnostic problem through the information base and the attributes which have been detected to date in the study of the patient. The inference function was based upon Bayesian model, and the current view held by the program was a conditional distribution for various diseases. The following schema was used for this:

$$P''(D_k/E'') = \frac{P(S_i/D_k, E')P'(D_k/E')}{\sum_j P(S_i/D_j, E')P'(D_j/E')}$$

where $P'(D_k/E')$ is the probability of D_k given the total experience to date,E', but before the observation of the attribute S_i, $P(S_i/D_k, E')$ is the conditional probability of attribute S_i given D_k and E', and $P''(D_k/E'')$ is the probability for D_k given the new, increased experience. Decisions regarding tests and termination were made on the basis of calculations of expected costs and benefits at each step. The performance of the system was tested in two problem areas, the diagnosis of bone tumors and the diagnosis of congenital heart disease. They used Warner's [32] probability matrix and a priori disease probabilities. The complete diagnosis employed by Warner and the sequential diagnosis applied by Gorry gave the same expected accuracy, but the latter system reached the final stage with an average of 6.9 tests because of the test selection function implemented in the program [1968].

Bleich and his associates [6] developed a program for estimation of acid-base disorders and then extended that program to consider electrolyte abnormalities. After the data collection has been performed, depending on the abnormalities of the data branched-chain, logic was activated and only the required sections of the decision pathways where analyzed. Questions asked during the process were either numerical laboratory values or "yes-no" type questions. Then, depending on the case analyzed, the program generated an evaluation note including suggestions regarding possible causes of the observed abnormalities and suggestions for correcting them. For this program there was no feedback: that is, the system was not referring to prior analysis of a patient, and every case was treated as a new one. [1969]

With the experience of their previous work, Gorry and his associates described the use of the discipline of decision analysis as the basis for an experimental interactive computer program designed to assist the physician in the clinical management of acute oliguric renal failure [15]. Their program was divided into two parts: phase I considered only tests with the minimal risk (e.g., historical data, chemical tests); and phase II involved tests of more risks and inconvenience. Phase I used

a sequential test selection process based on Bayes' Theorem, [14]. In phase II the methodology of decision theory was applied, where at each step in the "decision process" the program considered whether it was best to treat the patient immediately or to carry out additional diagnostic tests. The treatment with the current highest expected value was chosen and then compared with the expected values of the treatments that could be given if another diagnostic test were performed. The relevant values and probabilities of outcomes of treatment were obtained as subjective estimates from nephrologists. 18 test cases were evaluated and for 14 of the cases. the program selected the same therapeutic plan or diagnostic test as the expert. In three of the remaining four, the program's choice was expert's second choice [1973].

MYCIN developed by E. H. Shortliffe is a symbolic reasoning program and is considered state of the art among programs developed in this field. It determines the site of infection, type of organism and drug sensitivities of the organism which is used for antimicrobial therapy [27]. MYCIN has two kinds of data, the first is the patient data (the information about the patient which is entered in response to computer generated questions during the consultation). The other is the "dynamic data," which is a data structure created during the consultation. The program has three subcomponents: first a consultation system in which questions are asked and through which conclusions are drawn and advice is given. The second is an explanation system which answers questions from the user and attempts to explain its advice. The third is the rule-acquisition system which permits experts to teach MYCIN new decision rules and/or to alter pre-existing rules that are judged to be inadequate or incorrect. Decision in MYCIN, not only involves the patient but also the cultures that have grown, the organisms isolated, and drugs that have been administered. Each of these are termed a 'context' of the program's reasoning. The context-types initiated during a run of the program, are arranged hierarchically

in a data structure termed 'context tree' in which each node is represented as one 'context'. Rules are subject to categorization in accordance with the context types. Every rule in the system belongs to one and only one of these categories. A clinical parameter is a characteristic of one of the contexts in the context tree,i.e., the name of the patient, the site of a culture, and so on. MYCIN stores inferences and data using the attribute–object–value concept. *Object* is always some context in the context tree, and *attribute* is a clinical parameter appropriate for that context. The *value* of every clinical parameter is stored by MYCIN along with an associated certainty factor (CF) that reflects the system's belief that the value is correct. In addition each rule in MYCIN is assigned a certainty factor. The CF approach is used because in most of the cases clinicians do not use the "information comparable to implanted standard statistical methods". Certainty factors allow the accumulation of evidence and facilitate decisions concerning the identification of organisms causing diseases in patients. MYCIN's consultation session creates the patient context as the top node in the context tree. MYCIN then attempts to apply the goal–rule to the newly created patient context. The goal oriented approach to rule innovation and question selection is automated via two interrelated procedures, one is the rule analysis and the second is a mechanism that searches for the data needed by the first procedure [1976].

A system that was used for multiple disorders was developed by Ben-Bassat and his coworkers [4]. The knowledge base of the system consists of disorder patterns in a hierarchical way that was used as a feedback for medical information required for diagnosis. The system model consists of elements that includes 'features' and 'disorders'. Features were defined as bits of clinical data like age, sex, symptoms and others. For each feature a cost was assigned and shown in five ranks, like historical information and findings of physical examination were assigned as cost-1, inexpensive routine procedures were assigned as cost-2, and so on. Disorder was

defined as a feature or combination of features that describes a well-defined clinical entity like a problem or a disease. They were defined by means of a characterizing pattern which was composed of a set of features and conditional probabilities of features for a given disorder being present or absent. Rank of life threatening severity, and prior probability of appearance in the population under consideration, were also used to describe disorders. Knowledge extraction was done from disorder domain to the feature domain disorder characterization, disorders differentiation and feature characterization stages were applied until a high quality of pattern was recorded. In the diagnostic model analysis, each disorder was taken into account along with its complement, and each was individually considered. At any level of the program, the user might control the operation strategy and take the full control. The system is capable of providing reasons for it's decisions. The knowledge base of the system was incomplete and inaccurate [1980].

INTERNIST is a consultation program developed for internal medicine by H. Pople and his associates [24]. Its knowledge base is composed of disease entities and manifestations(symptoms, physical signs, and laboratory data). Each manifestation of a given disease is assigned two numbers: 'evoking strength' and 'frequency', with the values of 0–5 and 1–5 respectively, and 'import' is assigned for each manifestations across all disease with the values 1–5. There are two heuristic principles: one is the formation of problem areas through a partitioning algorithm and the other is the conclusion (or diagnosis) within a problem areas. During the diagnostic consultation the following steps are applied. First the positive and negative findings of the patient are entered by the user, for each positive manifestation given, the program retrieves its complete differential diagnosis. A disease hypothesis with a proper listing is created, and for each disease hypothesis four lists are maintained and each hypothesis on the master list of diagnoses are given a score. After the scoring, the master list of all hypotheses are sorted by descending score. The diag-

nosis whose score falls short of the threshold are discarded. Possible diseases for the likeliest diagnoses are identified from the master differential list by a partitioning rule. After selecting the most attractive diagnosis the step for a definitive diagnosis is applied. If there is no conclusion, the program either pursues,rules out or discriminates with a certain strategy. To improve the efficiency, the system asks questions and the program reruns again to find a new differential diagnosis. The program stops when the import value of 2 or less is observed. The program cannot analyze the multisystem problems. the database structure limits the program's ability to reason anatomically and temporally, and it can not recognize the subcomponents of an illness. On the basis of the deficiencies mentioned above, the same group developed CADUCEUS and they defined the diagnostic complex(es) from the beginning and applied facets of disease to more than one diagnostic entity to overcome the deficiencies. The authors believe that CADUCEUS will not be ready for release for another five to ten years [1982].

K. P. Adlassnig and his team has developed a data–driven,rule–based expert system for general medicine. called CADIAG [1] [2]. Using symbolic logic representation CADIAG-I was developed.then with some changes they developed CADIAG-II. The first version of CADIAG , CADIAG–I is based on a symbolic logic representation of a medical relationship. It consists of four main structures, namely a medical information system, a patient data interpreter, a computer assisted medical consultation system, and a medical diagnostic knowledge system. Medical entities such as; 1- symptoms, signs, laboratory findings, 2-diseases, diagnoses, 3- intermediate combinations, and 4- symptom combinations and their relationships are defined and represented in terms of first-order predicate calculus. Diagnostic hypotheses are generated by precalculating unique symptom patterns. For a given symptom pattern, a confirmed or excluded diagnosis, diagnostic hypotheses, and possible diagnoses are established. Diagnostic hypotheses are calculated by means of unique

symptom patterns matching the symptoms observed on the patients. Possible diagnoses are made on the basis of preferential symptoms exhibited by the patient and selected as such by the diagnostician. Extended explanations of the diagnostic results are given to the physician. Suggestions whether to examine the patient further in order to confirm or exclude diagnostic hypotheses or possible diagnoses are also offered [1985].

CADIAG-II is an expansion of CADIAG-I and uses Fuzzy Set Theory to determine the relations between symptoms and diseases. In the system's knowledge base symptoms,diseases or diagnoses and intermediate and symptom combinations are given some fuzzy logical values. To calculate the grades of the membership of the patient to disease, compositional rules of inference are used. The relationships between symptoms and diseases are described either linguistically or statistically. Symptoms are not present or absent only, but they are assigned a value between 0–1 to indicate the 'degree of membership'. Diseases and diagnoses are treated in a similar way. In the diagnostic process, after the symptoms are gathered, possible intermediate combinations and symptom combinations are computed. Contradictions in the present symptom pattern and the intermediate computed patterns of symptom combinations are checked. Then confirmed diagnoses are identified and diagnostic hypotheses are offered [1986].

Ohmann and his friends studied the extensions of the independent Bayes Model, taking interactions between variables into account, together with the data set of upper Gastrointestinal bleeding, using different measures of performance, such as discriminant ability, sharpness of prediction and reliability of the probabilities [22]. The models used were, linear logistic regression and independence Bayes. It was shown that there were small differences between the models if applied to data sets with few variables. With the data sets of many variables, there were sizable differences between the models, but no model was superior in all aspects of performance

[1988].

In his recent paper Adlassnig presented the performance evaluation of diagnostic accuracy of the medical expert system [3]. Taking histologically or clinically confirmed diagnosis as standard he showed that the ROC–Receiver Operating Characteristic–curves not only allow the optimal adjustment of the expert system's internal and hoc decision criteria such as thresholds, weights and scores but also provide a basis for better comparing the performance of different medical expert systems [1989].

Chapter 2

A BAYESIAN FORMULATION FOR CALOND

The purpose of this chapter is to estimate posterior lesion probabilities for a given set of neurological test outcomes using a classical statistical process: Bayes' Theorem.

As mentioned before in Section 1.1 the system CALOND will be designed to relate clinical test outcomes to the spatial distribution of the probability of lesions. For this purpose the anatomical structures of interest will be divided into V volume units, called *voxels*. Each voxel will be identified by some suitable code, designated here by the symbol v, $v = 1, 2, \ldots, V$. A neural pathway can then be represented as a string of voxels through which it passes.

One of the important tasks of the study will be the preparation of an elaborate database. The database will contain detailed anatomical information, such as relationships between anatomical structures of importance and voxels, a list of clinical tests and their different possible outcomes, and statistical information such as á priori probabilities of malfunction and function and various test outcomes. For practical purposes CALOND's database will be constructed from the following complementary prime units:

- TESTBASE : This will be CALOND's database unit in which the test inputs and the resulting test outcomes will be stored. The list of the tests and

outcomes for the CALOND is shown in Appendix D.

- PATHBASE : This will be the unit in which the pathway information will be stored, that is, the voxels which compose the pathway.
- ANATBASE: This will be CALOND's database unit used for storing the anatomical information, that is the anatomical names related to pathways is stored in this unit.
- STATBASE: This is the unit in which statistical information will be maintained. (Á prior information and conditional á prior information).

The ultimate objective of the study is to develop and implement a method for calculating the probability of lesion for each voxel (or pathway) based on the test outcomes. Such a method is introduced in the following subsection.

2.1 A Statistical Method for Localization of Lesions

Consider the following scenario : A person with (possible) neurological complications , henceforth to be called simply the *PATIENT* walks into a general practitioner's office. He, or she, is about to be examined by a medical professional who is using CALOND. This medical professional might be a physician, a physician's assistant, a medical technician or other qualified personnel. For the sake of conciseness, in this study, from now on this person will be referred to as the *PHYSICIAN*.

After an initial interview with the patient the physician is expected to apply a group of tests [1] from CALOND's database unit TESTBASE. If the patient has any lesions involving the nervous system, a skilled neurologist or a neurosurgeon can give a good estimate of the location of lesions by associating his or her experience with the test outcomes. In this section an analytical method will be developed for

[1] CALOND will not require the tests to be applied in any particular order.

the determination of the probability of malfunction of a given voxel (or pathway) for a given set of test outcomes. In the following, some pertinent terminology will be introduced and the problem statement and related conditional probability equations will be presented.

Let T_i $(i = 1, 2, \ldots, I)$ denote the verbal description of an *available test* in TESTBASE, $\mathcal{T} = \{T_i\} = \{T_1, T_2, \ldots, T_I\}$ *the set of available tests* and let

$\mathcal{O} = \{O_{1,1}, O_{1,2}, \ldots, O_{1,J(1)}; O_{2,1}, O_{2,2}, \ldots . O_{2,J(2)}; \ldots, O_{I,1}, O_{I,2}, \ldots, O_{I,J(I)}\}$

denote the *set of all test outcomes* in TESTBASE, where $O_{i,j}$ is the jth outcome of test i.

Note that these definitions imply that there are exactly $J = \sum_{i=1}^{I} J(i)$ outcomes in TESTBASE.

Let $O_i = \{O_{i,1}, O_{i,2}, \ldots, O_{i,J(i)}\}$ denote the set of possible outcomes for the ith test T_i, and let n be the test sequence applied to a specific patient where $n = 1, \ldots, N$. For each test applied, the test outcome will be represented by a vector $\boldsymbol{q}(n)$, which will include the identification number of the test applied and its test outcome from the data base, namely TESTBASE. Note that the physician need not follow any predetermined test sequence dictated by CALOND.

Let M_v represent *the event* that voxel v (or pathway v) is malfunctioning.

Using the definition of conditional probability, the probability of the voxel v malfunctioning given that the test outcome of the applied test $T_{q(n)}$ is $\boldsymbol{q}(n)$ can be determined in terms of the following á priori probabilities:[2]

1. $P(M_v)$: the á priori probability of voxel v being malfunctioning.

2. $P(q(n))$: the á priori probability of test outcome $\boldsymbol{q}(n)$ being observed.

3. $P(q(n)/M_v)$: The conditional probability of test outcome $\boldsymbol{q}(n)$ being observed given that voxel v is malfunctioning.

After the application of the first test, $(n = 1)$, the malfunction probability of

[2]These probabilities are for a population at large, visiting general physician's office for a medical examination.

voxel v for a given test outcome $q(1)$ is given by

$$P(M_v/q(1)) = P(q(1)/M_v)\frac{P(M_v)}{P(q(1))} \tag{2.1}$$

After the second test (n=2) the probability of malfunction of voxel v, given the test outcomes $q(1)$ and $q(2)$ can be calculated from;

$$P(M_v/q(2)q(1)) = P(q(2)/q(1)M_v)P(q(1)/M_v)\frac{P(M_v)}{P(q(2)q(1))} \tag{2.2}$$

For the rest of this analysis the following fundamental assumption will be made:

Assumption 1 :

Let T_v denote the set of tests designed specifically for testing the voxel v. Let $S_v = \{q_1, q_2, \ldots, q_\nu\}$ denote the corresponding set of test outcomes for the patient under consideration and M_v denote the event that voxel v is malfunctioning. Let $S_k \subset S_v$ and $S_k' \doteq S_v - S_k$, then

$$P\{S_k/S_k' \cap M_v\} = P\{S_k/M_v\} \tag{2.3}$$

Let $\{q(r), q(r-1), \ldots, q(1)\} \in S_v$ and let $S_k = \{q(r)\}$, then *Assumption 1* implies that

$$P(q(r)/q(r-1)q(r-2)\ldots q(1)M_v) = P(q(r)/M_v) \tag{2.4}$$

where $\{q(r), q(r-1), q(r-2)\ldots, q(1)\}$ is a set of outcomes corresponding to a set of tests designed **specifically** for testing the voxel (or, the pathway) v.

The above assumption simply states that in computing the conditional probability of the test outcome $q(r)$, given the previous test outcomes $S_{r-1} = \{q(r-1)q(r-2)\ldots q(1)\}$ and M_v (i.e., voxel v is not functioning), the knowledge of S_{r-1} may safely be discarded in view of the definitive knowledge of M_v.

As an example, let $r = 2$ and $v = 10$. Assume that $q(1)$ is a test outcome indicating that voxel 10 is malfunctioning with some probability. Clearly in the computation of the conditional probability of the outcome of a new test, when the

test outcome $q(1)$ is given and it is known that the voxel M_{10} is malfunctioning, $q(1)$ may safely be dismissed in view of the more decisive knowledge M_{10} (i.e., voxel 10 is malfunctioning). This is exactly what *Assumption 1* implies. Note that for $r = 2$, *Assumption 1* yields $P(q(2)/q(1)M_v) \approx P(q(2)/M_v)$, then the Eq. (2.2) can be written as follows;

$$P(M_v/q(2)q(1)) \approx P(q(2)/M_v)P(q(1)/M_v)\frac{P(M_v)}{P(q(2)q(1))} \tag{2.5}$$

After the third test, $T_{q(3)}$, is applied the probability of malfunction of voxel v, given the test outcomes $q(1)$ $q(2)$ and $q(3)$, can be calculated from the following equation:

$$\begin{aligned} P(M_v/q(3)q(2)q(1)) &= P(q(3)/q(1)q(2)M_v)P(q(2)/q(1)M_v) \\ &\quad P(q(1)/M_v)\frac{P(M_v)}{P(q(3)q(2)q(1))} \end{aligned} \tag{2.6}$$

Hence, using *Assumption 1* the following expression will be written;

$$\begin{aligned} P(M_v/q(3)q(2)q(1)) &\approx P(q(3)/M_v)P(q(2)/M_v)P(q(1)/M_v) \\ &\quad \frac{P(M_v)}{P(q(3)q(2)q(1))} \end{aligned} \tag{2.7}$$

Using mathematical induction the following fundamental equation can be obtained:

$$\begin{aligned} P(M_v/q(1)q(2)\ldots q(n)) &= P(q(n)/M_v)\ldots P(q(2)/M_v)P(q(1)/M_v) \\ &\quad \ldots\frac{P(M_v)}{P(q(1)q(2)\ldots q(n))} \end{aligned} \tag{2.8}$$

The above equation allows recursive update of the malfunction probabilities after the arrival of each new test outcome and therefore it will be called *The Malfunction Probability Update Equation (MPUE).*

Although the MPU Equation developed above gives the malfunction probabilities of voxels only, the same equation is also applicable for entire pathways. In fact, in practical applications it might be computationally more efficient to update the malfunction probabilities along a given pathway rather than individual voxels.

2.2 Determination Of Subsequent Testing

After CALOND evaluates the outcome of a new test using the MPUE, it can suggest the next test to be applied by first selecting the voxel(s) (or pathway) with the highest malfunction probability and then searches the test(s) associated with the same voxel(s) from its TESTBASE.

As an example, assume that the test:

" *Observe vocal cords during phonation*" is applied and the outcome

" *Left cord weak or paralyzed*" is observed.

In TESTBASE this test outcome can be found to be associated with voxel numbers 1017, 928, ...,168. Suppose that after the malfunction probabilities of these voxels are updated using MPUE, CALOND will determine that the voxel 538 will have the highest malfunction probability. It will then suggest the tests 2,3,5,10,14,16,17,18,20,24,49,55,57, and 59, since according to TESTBASE these tests are listed as being associated with voxel 538.[3]

But there remains a problem: that is, how do we arrive at the malfunction probability values for each voxel. In the current study, only the brainstem is considered, comprised by 1000 voxels as mentioned before. If the anatomical structure is extended beyond the brainstem, the number of voxels will increase, eventually making the need for probability values cumbersome. In the following chapter, we propose a method for estimating malfunction probabilities, assuming the appropriate data acquisition is possible.

[3]Refer to the Appendix E for test outcome designation

Chapter 3

A REGRESSION ANALYSIS MODEL

3.1 Introduction

Our goal, once again, is to estimate posterior probabilities of lesion based upon á priori knowledge and the results of neurological tests. This estimation will now take the following direction:

Bayes' Theorem identifies a classical statistical relationship between priori knowledge and the attainment of more conclusive lesion probability statements based upon neurological test outcomes. Since we may consider the posterior probability of lesion to be a function of the á priori probability of certain neurological test outcomes, we may estimate this function through a polynomial (which corresponds to a Taylor series expansion of this function)–at first, a first order one which is essentially a linear approximation.

Linear Regression Analysis introduced in this chapter is a plausible method, since it exhibits the linearity between posterior and á priori probabilities in Bayes' Theorem. By modelling this linear relation between the á priori and posterior probabilities via statistical predictor techniques it is then possible to obtain some of the necessary probability estimates. A concise definition of Regression Analysis

is elucidated in Appendix B.

In the previous chapter, the following equation was derived by modifying Bayes' theorem:

$$P(M_v/q(n)) = P(q(n)/M_v)\frac{P(M_v)}{P(q(n))} \tag{3.1}$$

where: $q(n)$ is the test outcome vector chosen for the nth test applied, M_v is the event that voxel v is malfunctioning (or there is a lesion at that voxel). The probabilities are defined as follows: $P(q(n)/M_v)$ is the conditional probability of choosing test outcome $q(n)$ given that voxel v is malfunctioning. $P(M_v)$ is the á priori probability of voxel v being malfunctioning. $P(q(n))$ is the á priori probability of observing that specific test outcome for the nth test applied. $P(M_v/q(n))$ is the probability of having a lesion at voxel v given the test outcome $q(n)$ for the nth test applied. The universe, in the context of this work, does not cover the population at large but only those people who are seeking a neurological evaluation. It is also assumed that some or all of the á priori probabilities for the CALOND data file called STATBASE are provided by a designated team of neurologists and radiologists. After the first test is applied, the Eq. 3.1 will be as follows:

$$P(M_v/q(1)) = P(q(1)/M_v)\frac{P(M_v)}{P(q(1))} \tag{3.2}$$

which, in turn sets up a table such as that shown in the example given in Table 3.1., where 1062.....326 indicate the malfunctioning voxels involved for the observed test outcome $q(1)$.

3.2 Regression Analysis

Using regression analysis one can estimate a linear or nonlinear relationship between the variables. Of course, alternative models yielding similar estimates are possible. By imposing the regression analysis criterion to the designated model, it is possible to obtain the estimated regression coefficients; the model will then provide prob-

Voxel v	$P(M_v)$	$P(q(1))$	$P(q(1)/M_v)$	$P(M_v/q(1))$
1062	10^{-6}	10^{-2}	10^{-4}	10^{-8}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
326	10^{-6}	10^{-3}	10^{-3}	10^{-7}

Table 3.1: Example of probability assignments for malfunctioning voxels based on a single test.

ability of lesion estimates which may be compared to the numbers generated by participating physicians.

The notations used in the following will replicate the notations used in "Applied Regression Analysis" by N. R. Draper and H. Smith [9].

A linear, first order model with two predictor variables can be given as follows:

$$y = \beta_0 + \beta_1 z + \beta_2 x_1 \tag{3.3}$$

and similarly,

$$x_2 = \beta_0^* + \beta_1^* z + \beta_2^* x_3 \tag{3.4}$$

where $\beta_0, \beta_1, \beta_2, \beta_0^*, \beta_1^*$ and β_2^* are the unknown parameters of the model to be estimated, and z is an "experience factor",e.g., taken to be, for the moment, the years of experience of the physician who is estimating the á priori probabilities. Using the data available, the estimates of these parameters will be calculated. The estimates of y and x_2 which are, $\hat{y}$ and $\hat{x}_2$ respectively, for the given values of x_1, x_3 and z will be denoted as follows:

$$\hat{y} = C + Bz + Ax_1$$

$$\hat{x}_2 = C^* + B^* z + A^* x_3$$

where $\hat{y}, \hat{x}_2$ are the predicted values of y and x_2 respectively (for a given input value). C, B, A, C^*, B^* and A^* are the estimates of the parameters, β_i, β_i^* given

above. To estimate the parameters in the model, the *Least Squares* method will be applied, hence –See Fig. 3.1–the *Residual Sum of Squares–S–*, that is:

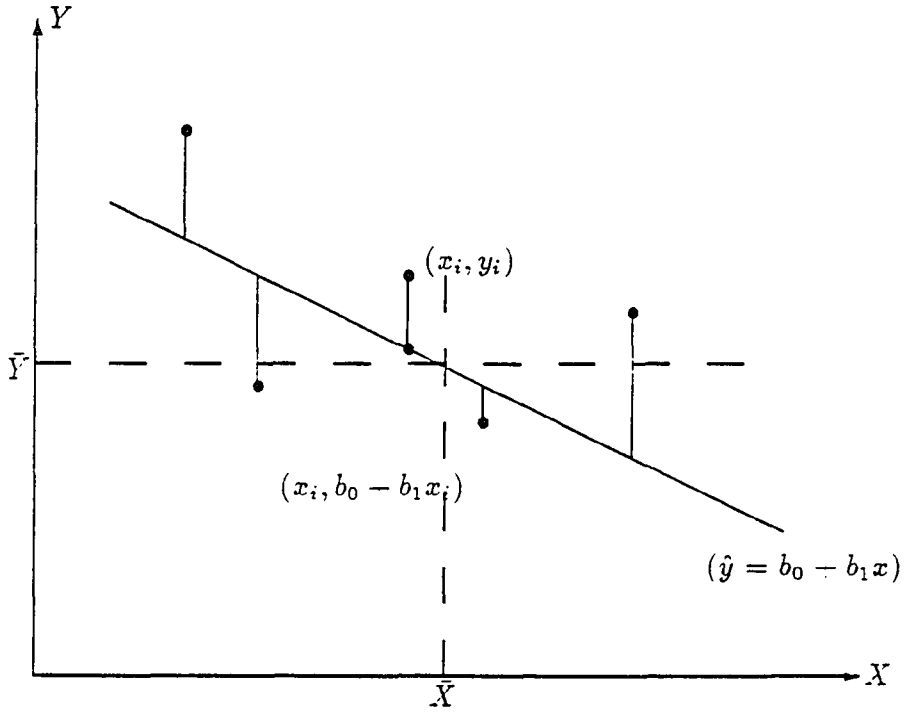


Figure 3.1: The vertical deviations whose sum of squares is minimized for the least squares procedure

$$S = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} (y_i - C - Bz_i - Ax_{1i})^2 \tag{3.5}$$

will be minimized. The least squares estimates of the parameters C, B, A in the model, found by minimizing the residual sum of the squares, are the solution to the following simultaneous system:

$$\frac{\partial S}{\partial C} = -2 \sum_{i=1}^{n} (y_i - C - Bz_i - Ax_{1i}) = 0$$

$$\frac{\partial S}{\partial B} = -2 \sum_{i=1}^{n} (y_i - C - Bz_i - Ax_{1i}) z_i = 0$$

$$\frac{\partial S}{\partial A} = -2\sum_{i=1}^{n}(y_i - C - Bz_i - Ax_{1i})x_{1i} = 0$$

These equations can be written as shown; they are referred to as the *Normal Equations*:

$$\begin{aligned} NC + (\sum z_i)B + (\sum x_{1i})A &= \sum y_i \\ (\sum z_i)C + (\sum z_i^2)B + (\sum x_{1i}z_i)A &= \sum y_i z_i \\ (\sum x_{1i})C + (\sum x_{1i}z_i)B + (\sum x_{1i}^2)A &= \sum y_i x_{1i} \end{aligned}$$

or in matrix form as follows:

$$\begin{bmatrix} N & \sum(z_i) & \sum(x_{1i}) \\ \sum(z_i) & \sum(z_i^2) & \sum(x_{1i}z_i) \\ \sum(x_{1i}) & \sum(x_{1i}z_i) & \sum(x_{1i}^2) \end{bmatrix} \begin{bmatrix} C \\ B \\ A \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i z_i \\ \sum y_i x_{1i} \end{bmatrix}$$

Applying Cramer's Rule, one can find the values of C, B and A. We may further observe that the system matrix

$$\begin{bmatrix} N & \sum z_i & \sum x_{1i} \\ \sum z_i & \sum z_i^2 & \sum x_{1i}z_i \\ \sum x_{1i} & \sum x_{1i}z_i & \sum x_{1i}^2 \end{bmatrix}$$

is guaranteed to possess an inverse since its determinant is strictly positive (this is a theorem from Regression Analysis: see [9] pp. 78–83)

Thus, our solution may be expressed as

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} N & \sum z_i & \sum x_{1i} \\ \sum z_i & \sum z_i^2 & \sum x_{1i}z_i \\ \sum x_{1i} & \sum x_{1i}z_i & \sum x_{1i}^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum y_i \\ \sum y_i z_i \\ \sum y_i x_{1i} \end{bmatrix}$$

And in fact,

$$\begin{bmatrix} N & \sum z_i & \sum x_{1i} \\ \sum z_i & \sum z_i^2 & \sum x_{1i}z_i \\ \sum x_{1i} & \sum x_{1i}z_i & \sum x_{1i}^2 \end{bmatrix}^{-1}$$

is known, in the literature, to be $V(b)/s^2$, where $V(b)$ is the variance, covariance matrix for the regression coefficient vector b; and s^2 is $SS_{Residual}/df_{Residual}$. In fact, the system matrix may be written in terms of variance and covariance. Since this matrix represents the first partial derivative of S, all of which are then known to be

positive, it assures us that the values of regression coefficients generated by Normal Equations do indeed identify a local minimum for S.

Now the *fundamental partition equation* is used to generate additional quantities in the ANOVA (ANalysis Of VAriance) table. Recall that [9]

$$\sum(y_i - \bar{y})^2 = \sum(\hat{y}_i - \bar{y})^2 - \sum(y_i - \hat{y}_i)^2 \tag{3.6}$$

is the fundamental partition equation, where; $\sum(y_i - \bar{y})^2$ is called *sum of squares total corrected* which is related to the variance of y, (σ_y^2). $\sum(\hat{y}_i - \bar{y})^2$ is called *sum of squares due to regression*, which shows how well data is regressed to y and $\sum(y_i - \hat{y})^2$ is *sum of squares due to error* (residuals) which serves as a measure of the inadequacy of the model. From this information it is possible to appraise whether the regression line will serve as a good predictor by observing how the Sum of Squares (SS) about the mean value of y has separated into the error SS or an explained (regressed) SS.

Next form the ANOVA (ANalysis Of VAriance) table as shown on the next page. In this table l is the sample size and k is the number of predictor variables. $T.C$ is called Total Corrected and it is equal to sum of the degrees of freedom due regression and the degrees of freedom due residual. The term MS is used to define the *mean square error* and it is equal to the ratio of SS and the degrees of freedom for a given category (source). The degrees of freedom can be found by subtracting the number of parameters in the model from the number of cases in the model. The term F-Ratio $= MS_{\text{reg}}/MS_{\text{res}}$ possesses an F distribution in this case.

Thus a test of the hypothesis $H_0 : B = 0$ or $A = 0$ versus $H_1 : B \neq 0$ and $A \neq 0$ can be conducted by examining the magnitude of the F-Ratio. If a certain level of significance (α) is defined, then we:

$$\text{Reject} \quad H_0 \quad : \quad \text{if} \quad F > F_{RC}$$

$$\text{Accept} \quad H_0 \quad : \quad \text{if} \quad F \leq F_{RC}$$

Source	Degrees of Freedom	SS	MS	F Ratio
Regression	$k\ (=2)$	$\sum(\hat{y}-\bar{y})^2$	$\sum(\hat{y}-\bar{y})^2/k$	
Residual	$l-3$	$\sum(y-\hat{y})^2$	$\sum(y-\hat{y})^2/(l-3)$	$F = MS_{\text{reg}}/MS_{\text{res}}$
T.C	$l-1$	$\sum(y-\bar{y})^2$	$\sum(y-\bar{y})^2/(l-1)$	

Table 3.2: The ANOVA Table (Used for SSE)

as shown:

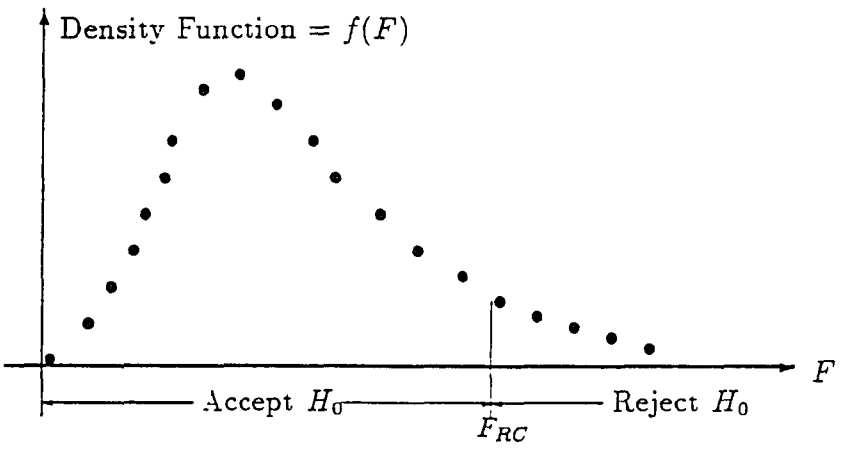


Figure 3.2: F-Distribution for α=significance level

The same method is applied for the estimation of $\hat{x}_2$. The following matrix is derived to find the estimates of C^*, B^*, A^*,

$$\begin{bmatrix} N & \sum(z_i) & \sum(x_{3i}) \\ \sum(z_i) & \sum(z_i^2) & \sum(x_{3i}z_i) \\ \sum(x_{3i}) & \sum(x_{3i}z_i) & \sum(x_{3i}^2) \end{bmatrix} \begin{bmatrix} C^* \\ B^* \\ A^* \end{bmatrix} = \begin{bmatrix} \sum x_{2i} \\ \sum x_{2i}z_i \\ \sum x_{2i}x_{3i} \end{bmatrix}$$

A similar ANOVA table will also be formed for this case.

3.3 Linear Hypothesis Testing

As an alternative to the model $y = \beta_0 + \beta_1 z + \beta_2 x_1$, consider the following model:

$$y = \beta_0 + \beta_2 x_1,$$

which discards the experience factor. In this case the regression model hypothesis will be $H_0 : \beta_2 = 0$ and the alternative $H_1 : \beta_2 \neq 0$.

For the reduced model, as before, least squares analysis is applied and the following matrix is obtained.

$$\begin{bmatrix} N & (\sum(x_i)) \\ \sum(x_i) & \sum(x_i^2) \end{bmatrix} \begin{bmatrix} b_0 \\ b_2 \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum y_i x_i \end{bmatrix}$$

Source	d.f	SS	MS	F Ratio
Regression	$k\ (=2)$	$\sum(\hat{y}-\bar{y})^2$	$\sum(\hat{y}-\bar{y})^2/k$	
Residual	$l-2$	$\sum(y-\hat{y})^2$	$\sum(y-\hat{y})^2/(l-2)$	$F = MS_{\text{reg}}/MS_{\text{res}}$
T.C	$l-1$	$\sum(y-\bar{y})^2$	$\sum(y-\bar{y})^2/(l-1)$	

Table 3.3: ANOVA Table for Reduced Model (Used for SSW)

If the term SS_{res} is computed for both models, the following analysis is possible (see Tables 3.1 and 3.2 respectively):

$$SS_{\text{res}} = \frac{\sum(y-\hat{y})^2}{l-3} = SSE \quad \text{for the expanded model}$$

$$SS_{\text{res}} = \frac{\sum(y-\hat{y})^2}{l-2} = SSW \quad \text{for the reduced model}$$

It is expected that $SSW > SSE$ since there are fewer parameters in the reduced model. Define $SSW - SSE$ as sum of squares due to errors incurred by the revision hypothesis $H_0 : C = 0$, which claims that the reduced model is superior: with only one degree of freedom: $|(l-2)-(l-3)| = 1$. Then, the validity of the revision hypothesis test will be tested by considering the following ratio of mean squares:

$$F = \frac{\frac{SSW-SSE}{1}}{\frac{SSE}{(l-3)}}$$

and then referring to the F distribution table with a given significance level for acceptance or rejection of H_0. Draper/Smith, pp. 102–107 [9].

Consider the modified Bayes' Equation that was derived for the system.

$$P(M_v/q(1)q(2)\cdots q(n)) = P(q(n)/P(q(1)q(2)\cdots q(n)M_v)\frac{P(M_v)}{P(q(1)q(2)\cdots q(n))}$$

After the application of the first test the above equation has the form:

$$P(M_v/q(1)) = P(q(1)/M_v)\frac{P(M_v)}{P(q(1))}$$

Let, $P(q(1)/M_v) = x_1, P(M_v) = x_2, P(q(1)) = x_3$, and $P(M_v/q(1)) = y$. Then the equation above will become

$$y = x_1\frac{x_2}{x_3}$$

or, in ratio form

$$\frac{y}{x_1} = \frac{x_2}{x_3} = a$$

where a is a probability ratio (proportionality) constant. It is possible, then, to formulate a pair of linear, first order regression models as shown below:

$$y = b + ax_1 \tag{3.7}$$

$$x_2 = c + ax_3 \tag{3.8}$$

which share this common probability ratio, a. After the application of the second test, the modified Bayes' Equation will be:

$$P(M_v/q(1)q(2)) = P(q(2)/q(1)M_v)P(q(1)/M_v)\frac{P(M_v)}{P(q(1)q(2))} \tag{3.9}$$

and with the use of Assumption 1, the Eq. 3.9 can be simplified to:

$$P(M_v/q(1)q(2)) = P(q(2)/M_v)P(q(1)/M_v)\frac{P(M_v)}{P(q(1)q(2))} \tag{3.10}$$

Now, let $P(M_v/q(1)q(2)) = y_1$, $P(q(2)/M_v) = x_5, P(q(1)q(2)) = x_4$; thus Eq. 3.10 can be written as:

$$y_1 = x_1x_5\frac{x_2}{x_4}$$

let

$$\frac{y_1}{x_5} = \frac{x_1x_2}{x_4} = a^*$$

then;

$$y_1 = b^* + a^*x_5$$

and

$$x_2 = c^* + a^*(\frac{x_4}{x_1}).$$

If we define $z = x_4/x_1$, then

$$x_2 = c^* + a^* z$$

Here the estimates of a^*, b^*, c^* will be calculated, which in turn will allow us to extrapolate x_4 using;

$$x_4 \approx \left[\frac{x_2 - c^*}{a^*}\right] x_1$$

In this fashion, it is possible to find the joint probabilities of observed test outcomes. Indeed this can be extended to estimate the general joint probability $P(q(1)q(2)\cdots q(n))$ for n observed test outcomes. This technique makes it possible to find many joint probability values, but requires a large database of á priori probabilities, currently unavailable. We then turn to available simulation strategies.

3.4 Conclusions

This chapter's primary focus is to use linear approximation techniques in order to estimate the posterior lesion probability for some anatomical region. The principal approach is to apply the classical statistics process of regression analysis. If all of the necessary á priori information is available, one can develop a storehouse of posterior lesion probabilities associated with a variety of neurological testoutcomes and general demographic statements (including epidemiological statements of incidence) for a host of neurological diseases. The physician would have access to this storehouse of information and then be able to use posterior lesion probability estimates to assist him in judgements as to how to proceed in the completion of the neurological examination and render his diagnosis and/or conclusions.

Chapter 4

AN ALTERNATIVE: MONTE CARLO SIMULATION

4.1 Definitions and Analysis

In most of the cases when the empirical data is not available, a randomized simulation may be applied to extract estimates for the values one seeks.

In Chapter 2, Bayes' theorem was introduced to find the probability of a malfunctioning voxel for a given set of test outcomes. In Chapter 3, it was discussed that á priori probabilities are needed to estimate the conditional probabilities of lesion. In this chapter, Monte Carlo Simulations are introduced in order to provide some information relating underlying probability–of–lesion–distributions of patients seeking neurological examinations and their subsequent test outcomes. The Monte Carlo Simulation is a commonly used technique for analyzing complex, statistical problems. A brief definition and relevant terminology about simulations is given in Appendix B.

As a first step, a portion of the anatomical structure of the nervous system, the brainstem, is subdivided into 20 subsections (left and right sides for 10 perpendicular to the long axis of the brainstem). This gives a total of 50 voxels for each subsection. The elementary part of the structure is the volume element, or voxel, v. For each

voxel v, some relevant factors will be introduced, mainly the Malfunction Factor $M = M(v)$ as the number of involvements of voxel v in pathways of probable malfunction and the Function Factor $F = F(v)$ as the number of involvements of voxel v in pathways of probable function, each pathway being associated with an observed test outcome. The Net–Malfunction Factor, $NMF = NMF(v)$ is defined as the difference between the Malfunction Factor (M) and Function Factor (F), i.e., $NMF = M - F$ for a given voxel or pathway, etc.. Therefore the Net-Malfunction Factor will appropriately include both the Malfunction and Function Factors in its structure and it is possible to say that the positive factors will indicate the locations of probable malfunction whereas the negative factors will indicate the locations of probable function. In this respect the Net-Malfunction Factor will have the following values.

$$NMF = \begin{cases} > 0 & \text{if } M > F \\ = 0 & \text{if } M = F \\ < 0 & \text{if } M < F \end{cases}$$

On the basis of these intervals, it is possible to assign the following parameters to a set of Net-Malfunction Factors in a given region s.

$$F_s = \left|\left[\sum \text{negative } NMF \text{ in } s\right]\right|, \text{and } M_s = \left[\sum \text{positive } NMF \text{ in } s\right].$$

Then it is possible to visualize a relation between F_s and M_s as shown: F_s is the horizontal coordinate; M_s is the vertical coordinate in a standard Cartesian plane. We then pose a function for comparison of M_s and F_s magnitudes as: $\tan\theta = \dfrac{M_s}{F_s}$ and $0 \leq \tan\theta \leq \infty$, or $0 \leq \theta \leq \pi/2$. There are some regions of interest, named as follows:

$0 \leq \theta < \theta_k$	Function Region
$\theta_k \leq \theta < \theta_l$	Ambiguous Region
$\theta_l \leq \theta \leq \pi/2$	Malfunction Region

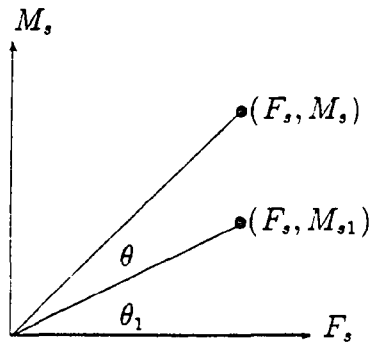


Figure 4.1: F_s and M_s to generate θ

The relation $M_s = (\tan\theta)F_s$ can be shown explicitly as follows:

$$(m_1 + m_2 + \cdots + m_m) = (\tan\theta)(f_1 + f_2 + \cdots + f_f)$$

(where the m_i's are the positive Net-Malfunction Factors in the current subsection, and the f_j's are negative Net-Malfunction Factors in the current subsection). If only the ith voxel is taken into account, then the relation between this voxel and the Malfunction Factors and Function Factors is as follows:

$$m_i = (\tan\theta)|(f_1 + f_2 + \cdots + f_f)| - \underbrace{(m_1 + m_2 + \cdots + m_m)}_{\text{except } m_i} \tag{4.1}$$

that is

$$m_i = (\tan\theta)F_s - M_{s1} \tag{4.2}$$

where $M_{s1} = \underbrace{(m_1 + m_2 + \cdots + m_m)}_{\text{except } m_i}$. From Fig. 4.1 above it can be seen that the relation between M_{s1} and F_s will be given as,

$$M_{s1} = (\tan\theta_1)F_s$$

therefore the Eq. 4.2 will be written as follows;

$$m_i = (\tan\theta)F_s - (\tan\theta_1)F_s$$

or

$$m_i = \Big(\tan\theta - \tan\theta_1\Big)F_s$$

and using trigonometric identities the equation can be written as follows;

$$m_i = \frac{\sin(\theta - \theta_1)}{\cos\theta\cos\theta_1}F_s \tag{4.3}$$

To see the behavior of this function, the first and the second derivatives of m_i will be taken with respect to θ: and θ_1 will be taken as a parameter. The result of the first derivative will yield the following;

$$\frac{dm_i}{d\theta} = F_s\sec^2\theta$$

which shows that it is positive and thus m_i is a monotone increasing function. The second derivative will be equal to

$$\frac{d^2m_i}{d\theta^2} = 2F_s\frac{\sin\theta}{\cos^2\theta}$$

which will be greater then zero as long as θ is in the first quadrant. It shows that the graph of m_i is concave up for $0 \leq \theta \leq 90°$ and is concave down for $90° \leq \theta \leq 180°$ and $90°$ will be the point of inflection for the graph of m_i as a function of θ. From Eq. 4.3 the following is derived.

$$\sin(\theta - \theta_1) = \frac{m_i}{F_s}\cos\theta\cos\theta_1$$

and

$$\cos\theta = \sqrt{\frac{1}{1 + (\frac{m_i}{F_s} + \tan\theta_1)^2}}.$$

The graph of Eq. 4.3 is shown in Figure 4.2

Referring to Fig. 4.1. it is also possible to write;

$$\frac{M_s}{I_s} = \sin\theta,$$

and,

$$\frac{F_s}{I_s} = \cos\theta$$

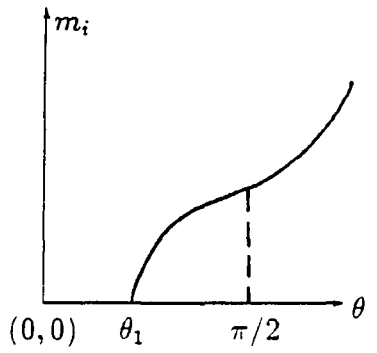


Figure 4.2: Relation between m_i versus θ. (Eq. 4.3)

thus.

$$\frac{d}{d\theta}\left[\frac{M_s}{I_s}\right] = \frac{F_s}{I_s}$$

which can be written as.

$$d\left[\frac{M_s}{I_s}\right] = \frac{F_s}{M_s}d\theta$$

or

$$d\left[\frac{F_s}{I_s}\right] = -\frac{M_s}{I_s}d\theta.$$

Assume that $d\theta \approx \theta - \theta_1$, then there is an alternative approach to estimate the change in M_s for a small change in θ and it is also possible to observe the relation between θ and θ_1 for the given values of m_i and F_s. For example, if $m_i = 0$, then $\cos\theta = \cos\theta_1$, the 45° diagonal on the θ versus θ_1 graph. Overall the relation between θ and θ_1 is used to map out the regions shown in Fig. 4.3. In this figure, M= Malfunction. A= Ambiguous and F= Function, are the regions that are defined above.

The Involvement Factor will be defined as $I_s = \sqrt{M_s^2 + F_s^2}$. Since the Involvement Factor is a measure of confidence in the estimates of θ and θ_1 it is expected that larger values of I_s should yield smaller values of $(\theta - \theta_1)$, which is indeed the case.

Representation of boundaries between ranges of malfunction, ambiguity and function, may be chosen in accordance with the magnitude of "Involvement" to

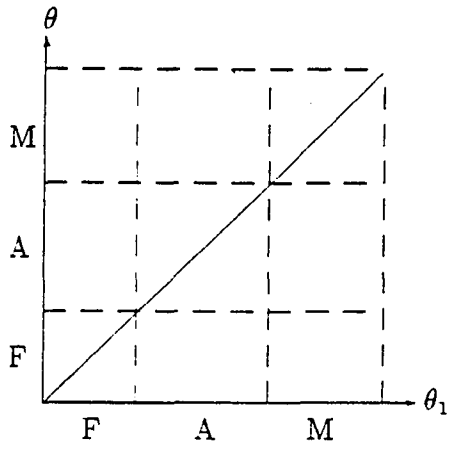


Figure 4.3: Regions of Function, Ambiguity, and Malfunction

reduce premature interpretations.

4.2 Simulation

For the actual simulation the following probability model is conjectured

$$P_s = c_s \cdot \frac{T_s \cdot M_s}{V_s \cdot I_s} \tag{4.4}$$

where c_s is a calibration factor, T_s is the total number of the tests applied in the subsection. V_s is the total number of voxels in the subsection. M_s is defined as the total number of positive Net–Malfunction factors and I_s is the Involvement factor, defined as $I_s = \sqrt{M_s^2 + F_s^2}$, where F_s is the total number of negative Net–Malfunction factors. The variable P_s denotes the probability of malfunction for the subsection. This model is chosen since it encapsulates the linear relationship between P_s and M_s. Furthermore, it allows the constant of proportionality to incorporate important attributes, such as the number of tests applied, the number of implicated voxels, and a measure of involvement. For each subsection the calibration factor will differ, always computed to generate a P_s function which satisfies all of the

axioms of probability. Thus it depends upon the á priori definition of the probability sample space and the density function. Knowing the probability distribution of the subsection it will then be possible to find the probabilities of malfunction for individual voxels.

4.2.1 Example

In this subsection, for given data–the information obtained from a variety of medical textbooks [16],[8]–about the distribution of lesions in the brainstem area for *Friedreich's Ataxia* will be presented and used to compute a theoretical calibration factor which will then serve for comparison with the results of a Monte Carlo simulation that focuses upon a typical neurological examination for a patient suffering from Friedreich's Ataxia. The distribution of lesions for *Motor Decussation* of the brainstem is shown in Fig. 4.4.

In the simulation process, the random numbers are generated to correspond to a normal density function, since in this case the population at large consists of patients with neurological diseases, specifically Friedreich's Ataxia.

Recalling the rule of "three sigmas", one can assert that for a normal density $p(x)$:

$$\int_{a-3\sigma}^{a-3\sigma} p(x)dx = 0.997 \tag{4.5}$$

In Eq. 4.5, a is the mean and σ is the standard deviation of the normal density $p(x)$. Let (a', b') be an arbitrary interval contained in $[a, b]$ (that is $a \leq a', b' \leq b$). The probability that a random variable X lies in the interval (a', b') is equal to the integral

$$P\Big(a' < X < b'\Big) = \int_{a'}^{b'} p(x)dx \tag{4.6}$$

Using this relation together with equation 4.5, one can say that

$$P\Big((a - 3\sigma) < X < (a + 3\sigma)\Big) = 0.997 \tag{4.7}$$

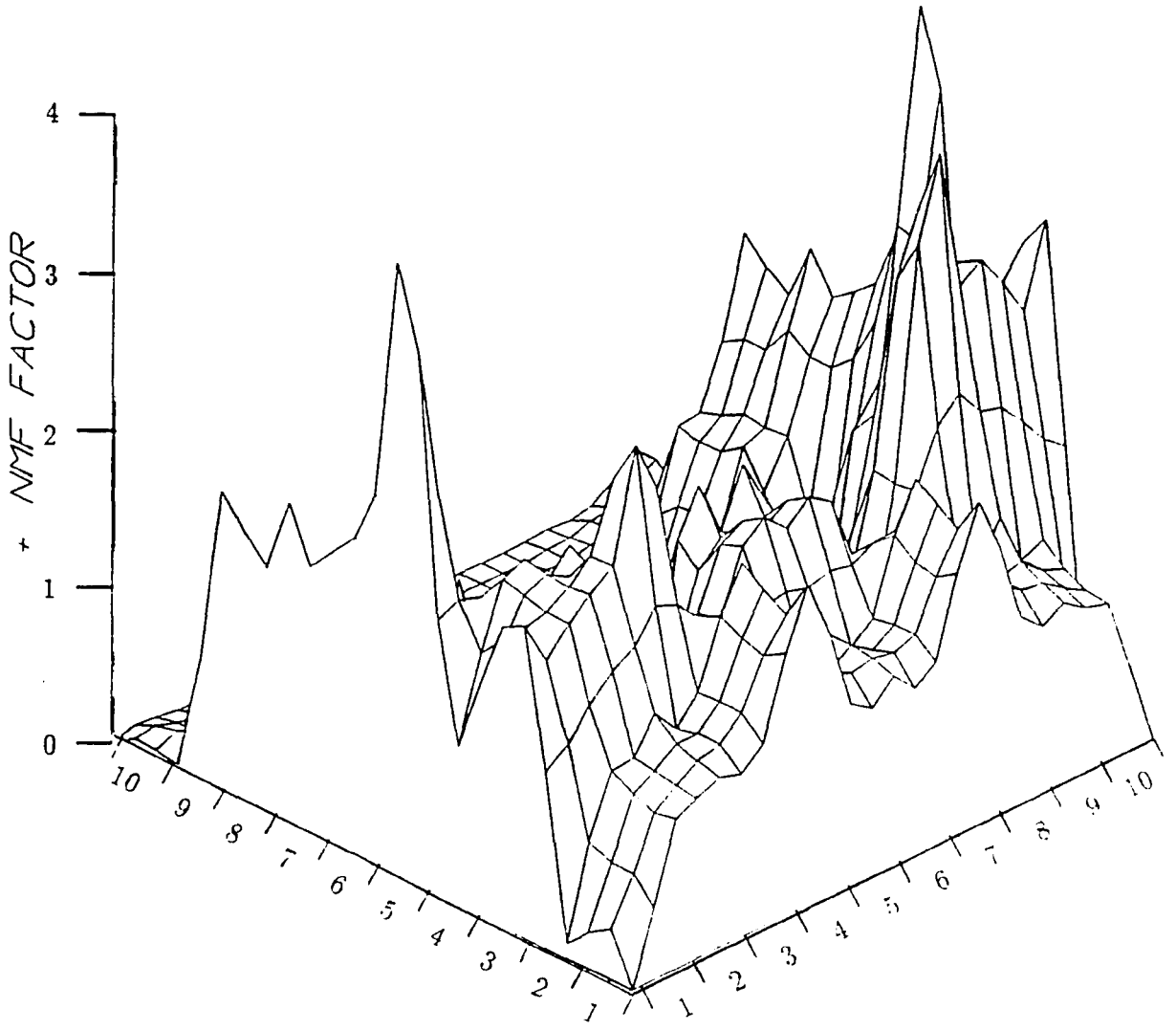


Figure 4.4: Distribution of positive net-malfunction factors in brainstem section through motor decussation for Friedreich's Ataxia

almost unity [29]. In the simulation, the mean of the normal density is zero and it is desired to capture an interval of length 6σ for a random number range of -1 to 1.

The normal density distribution with the corresponding random number intervals for the incrementation factors of Malfunctioning and Functioning are shown in Fig. 4.5.

For each random number generated, depending on which target of A, B, C, the random number hits, the malfunction factor of the voxels assumed to be malfunc-

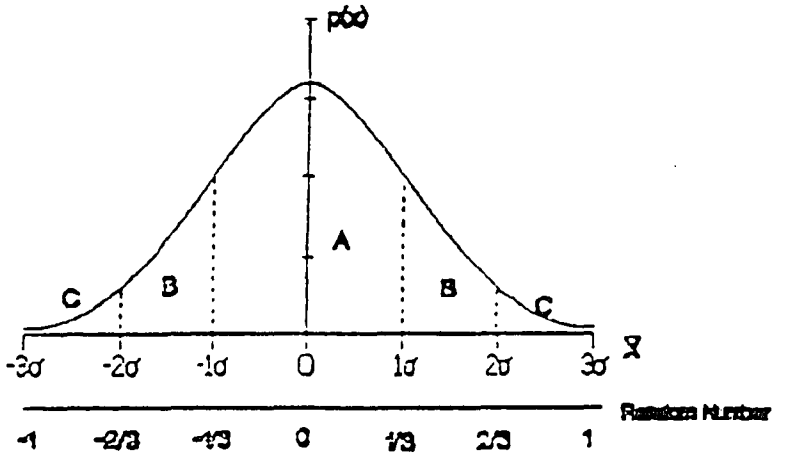


Figure 4.5: Normal Density Distribution with $a = 0$ and $\sigma = 1$

tioning are incremented by (0.67, 0.95, 0.97) and the function factor of the voxels assumed to be functioning are incremented by (0.33, 0.05, 0.03) respectively. This is done for the entire section for all the voxels implicated by test outcomes. Then it is possible to calculate the positive Net-Malfunction factor (M_s) and the negative Net-Malfunction factor (F_s), and the Involvement factor (I_s). With a given á priori probability for this specific section to malfunction, it will then be possible to find the calibration factor for the section. Fig. 4.6 is plotted from the results of the simulation for the Motor Decussation section of the brain stem.

These figures are generated by using the real and simulated calibration factors which are generated in the simulation. It is also possible to consider small subsections of the section under study, and to then estimate the probabilities of malfunction for these subsections. This will be discussed in the next chapter. Finally, the following table is presented to illustratively compare the calibration factor estimates for the real data and for the simulated data. It is possible to say that, if the brainstem section under consideration possesses a rather dense lesion distribution, then the values of the theoretical and simulated calibration factors will hardly differ; but if the lesion is diffused, this lack of geometrical connectivity creates a sometimes sizable gap in calibration factor estimation.

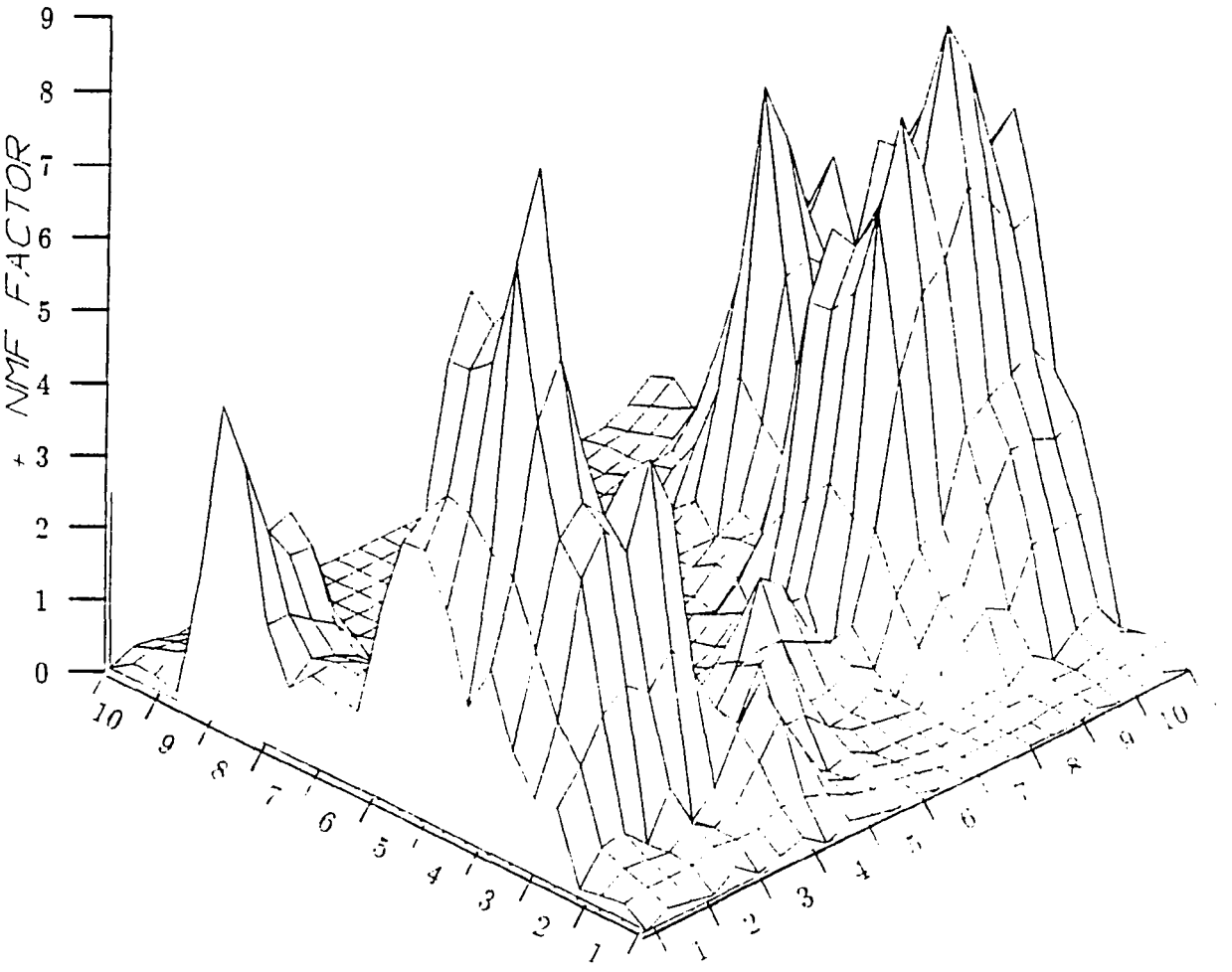


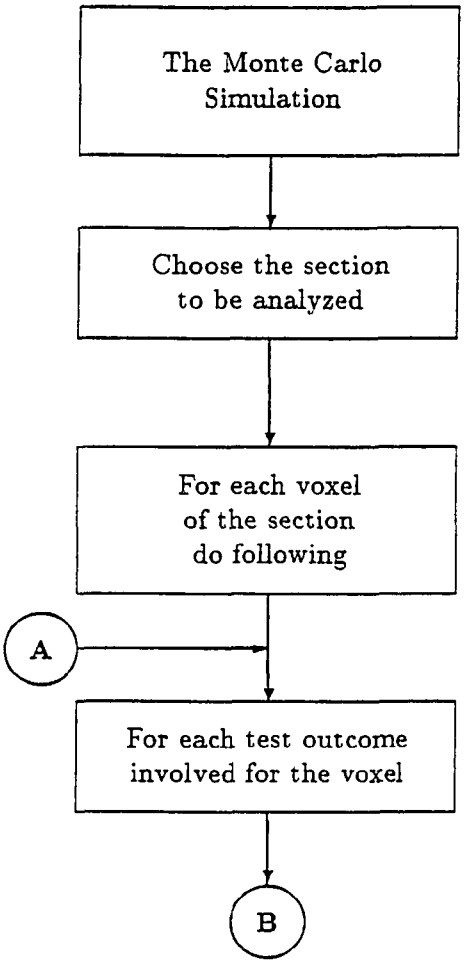
Figure 4.6: Simulated distribution of positive net–malfunction factors in brainstem section through motor decussation for Friedreich's Ataxia

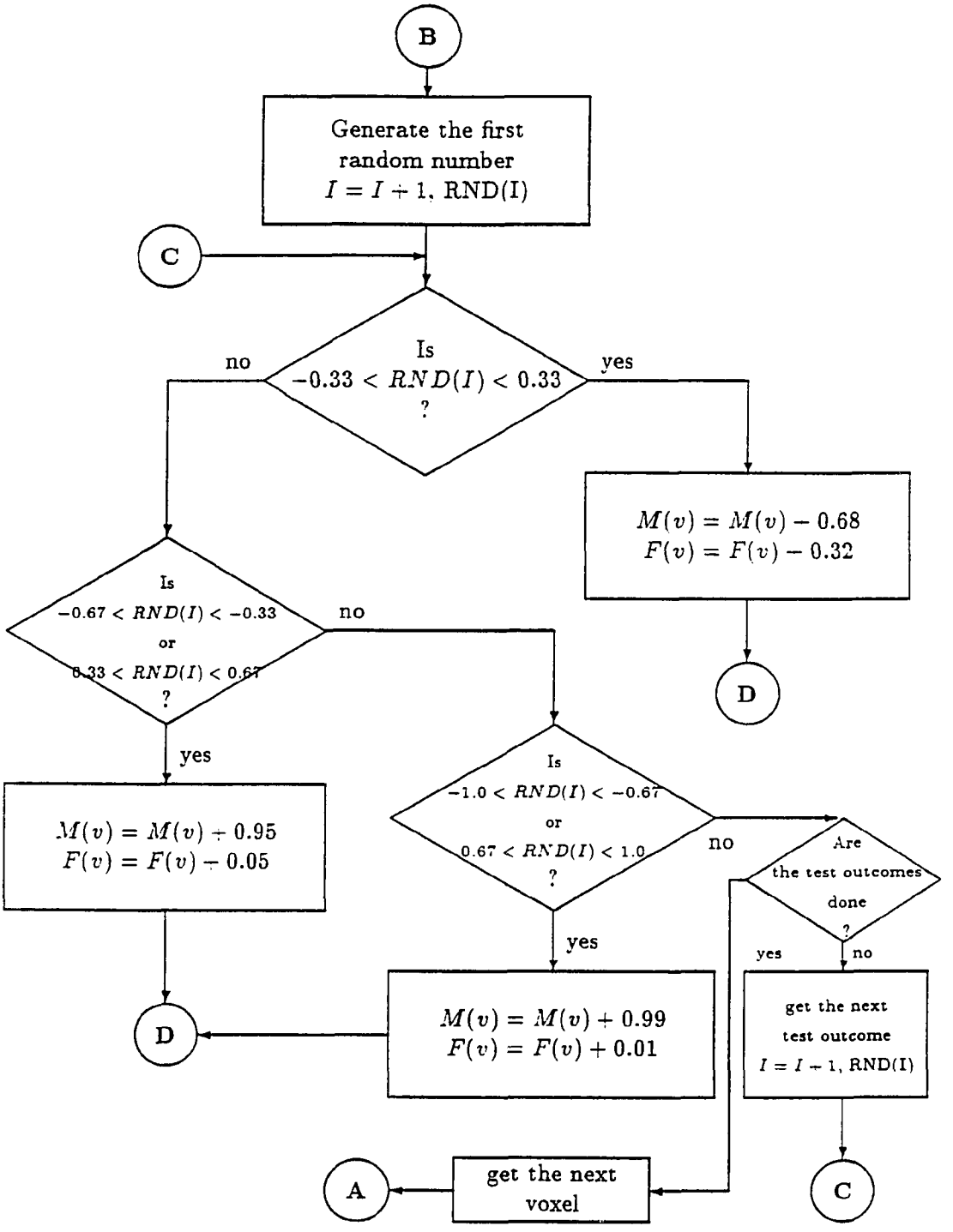
Sec.	Real Data-C_s	Simulation-C_s	% Error
1	3.3391582	3.39130	% 1.534
2	2.9230769	2.92307	% 0.0
3	1.6969697	1.69697	% 0.0
4	2.0121212	2.01212	% 0.0
5	1.6727272	1.67272	% 0.0
6	1.7043333	1.50000	% 11.98
7	0.9272727	0.92727	% 0.0
8	1.4716008	1.47027	% 0.09
9	1.7396416	1.62285	% 6.71
10	3.4244444	1.47368	% 56.96

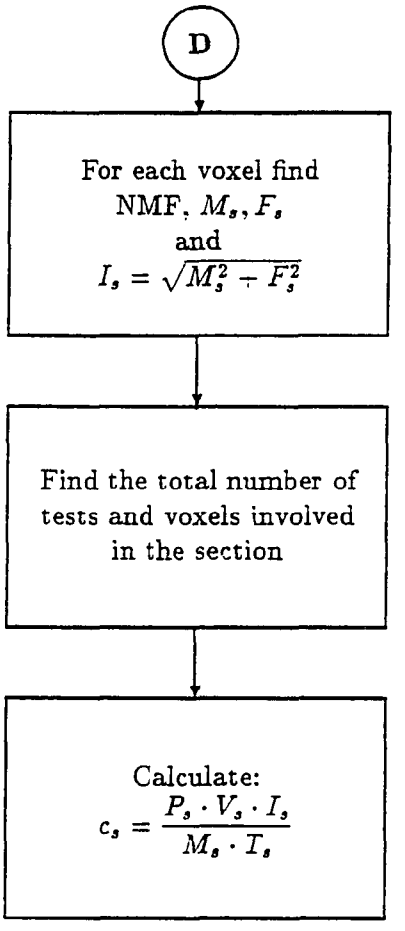
Table 4.1: Comparison of Real and Simulated Calibration Factors

The level of accuracy in this simulation is measured by computing a percent error. The extend to which the underlying probability distribution is disceretly approximated will effect the precision of the simulation. For increased accuracy, we may approximate the underlying continuos probability density by a discrete histogram with fragment intervals which are smaller. This in conjunction with an increase in the number of trials for the simulation would generate a desired level of accuracy.

Details of the simulation are available in the next chapter and the flowchart for the simulation is presented on the following pages.


The Monte Carlo Simulation
Choose the section to be analyzed
For each voxel of the section do following
A
For each test outcome involved for the voxel
B


B
Generate the first random number $I = I + 1$, RND(I)
C
Is $-0.33 < RND(I) < 0.33$?
no
yes
$M(v) = M(v) - 0.68$
$F(v) = F(v) - 0.32$
D
Is $-0.67 < RND(I) < -0.33$ or $0.33 < RND(I) < 0.67$?
no
yes
$M(v) = M(v) + 0.95$
$F(v) = F(v) + 0.05$
D
Is $-1.0 < RND(I) < -0.67$ or $0.67 < RND(I) < 1.0$?
no
yes
$M(v) = M(v) + 0.99$
$F(v) = F(v) + 0.01$
Are the test outcomes done ?
yes
no
get the next test outcome $I = I + 1$, RND(I)
C
get the next voxel
A


D
For each voxel find
NMF, M_s, F_s
and
$I_s = \sqrt{M_s^2 + F_s^2}$
Find the total number of
tests and voxels involved
in the section
Calculate:
$c_s = \frac{P_s \cdot V_s \cdot I_s}{M_s \cdot T_s}$

4.3 Conclusions

This chapter focuses upon simulation strategies to associate posterior probabilities of lesion, for a given anatomical region, with an underlying patient population possessing demographic and epidemiological attributes in conjunction with a predisposition to certain neurological test outcomes. This simulation lays the groundwork for the development of a library of posterior lesion probability statements which may be associated to populations which possess certain neurological diseases. In order to create this library, one must expand upon the neurological information base as well as develop techniques for lesion localization for diseases which are characterized by a diffused and disconnected set of lesions.

Chapter 5

DESCRIPTIVE STATISTICS FOR THE SIMULATION

Descriptive measures which indicate where the center or most typical value of a data set lies are called *Measures of Central Tendency*, often most simply referred to as averages [33]. In the following discussion, some measures of central tendency, (the mode, the median and the mean) and variation will be discussed in relation to the probability models posed for implementing CALOND system data in lesion probability estimation.

Our goal: Devise a method for "zeroing in" on troublesome subregions (i.e., locate regions with high probability of lesion–and focus upon the centers of such regions as crucial voxels for lesion localization).

5.1 Modal discussion of a linear probability density model

The **mode** of a data set is defined to be the data value or values that occur most frequently. A data set can have more than one mode. In our model, the mode is defined as the maximum of the sum of the positive net malfunction factors. The related calculations are shown below.

Consider the probability model of Eq. 4.4:

$$P_s = c_s \frac{T_s}{V_s} \frac{M_s}{\sqrt{M_s^2 + F_s^2}} \tag{5.1}$$

where $P_s = P\{Lesion\ in\ R_s/Q_s\}$ is the probability of having a lesion in region R_s for a given set of test outcomes–Q_s, c_s is the *calibration factor*, M_s is the sum of all the positive net malfunction factors in region R_s, F_s is the absolute sum of the all negative net malfunction factors involved in R_s. The Involvement factor is defined as $I_s = \sqrt{M_s^2 + F_s^2}$. T_s is defined as the total number of test outcomes involved in the region and V_s is the total number of voxels involved in the region. As an example, a section (R_s) with related net malfunction factors, is shown in Fig. 5.1.

P_s the initial probability of lesion, for R_s, is obtained as an estimate from a physician. Knowing P_s, it is possible to obtain the corresponding calibration factor: that is

$$c_s = \frac{P_s \cdot V_s \cdot I_s}{T_s \cdot M_s} \tag{5.2}$$

The discrete probability of malfunction for any subset of a fundamental set of voxels can be modeled as follows:

$$f(M) = kM \tag{5.3}$$

for that subset (which could be a voxel or a set of voxels). For the initial region the probability of Malfunction is $f_s(M) = k_s M$ where $0 \leq M \leq M_s$; and with the initial condition $f_s(M_s) = k_s M_s$, where k_s is defined as the *inverse malfunction constant.*

In this model $f_s = P_s$ therefore $k_s = P_s/M_s$. A linear probability function model is plausible because of the direct proportionality relation which exists between the sum of the positive Net-Malfunction Factors and the overall probability of lesion for a given subset of voxels under consideration.

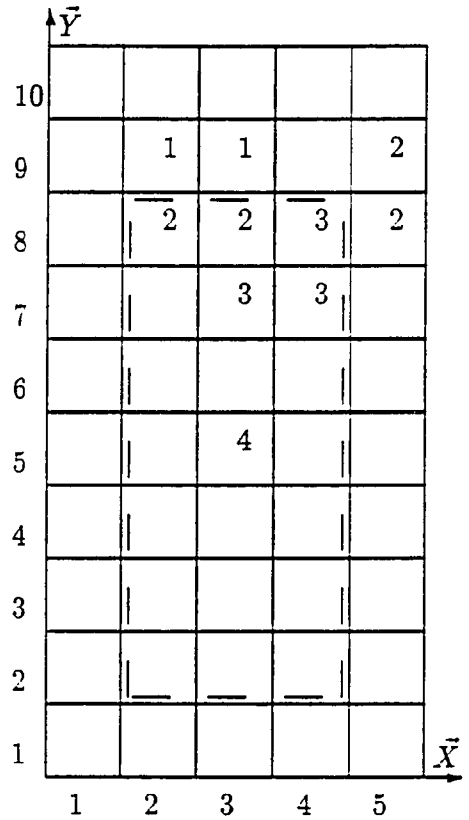


Figure 5.1: An example of a section R_s with a distribution of net malfunction factors. Enclosed with dashed lines is a subregion R_1, also shown in Fig. 5.2

Consider the subregion R_1, that is $R_1 \subset R_s$. The probability of having a lesion in R_1 given the same set of test outcomes as in R_s will be defined as,

$$P\{Lesion\ in\ R_1/Q_s\} = P_1 = k_1 M_1 \tag{5.4}$$

or

$$P_1 = c_1 \frac{T_1}{V_1} \frac{M_1}{\sqrt{M_1^2 + F_1^2}} \tag{5.5}$$

Assuming, the same sequence of random numbers, but excluding tests that do not impact upon R_1, one can make the following conditional probability statement;

$$P\{Lesion\ in\ R_1/Q_s\} = P\{Lesion\ in\ R_1/Q_1\ in\ R_1\} \tag{5.6}$$

What this will mean is that $P_1 = k_s M_1$, hence, $k_s = k_1 = \cdots = k_n$. Therefore as new subregions are taken into account the new calibration factor may be calculated as follows;

$$c_n = \frac{k_n \cdot V_n \cdot I_n}{T_n} \tag{5.7}$$

where n is the current identification number of the subregion under study. In general, it is possible to say that the slope of the linear density function remains constant and what changes is the calibration factor for each subregion under analysis.

Let us display the computer simulated results for the above process–in the figures shown, the Net Malfunction Factors with 0 values are not indicated even though they were counted as entities for the voxels taken into account–. Assuming that the initial probability of lesion within the whole space is $P_s = 1$, and identifying the simulation results as: $M_s = 23, F_s = 0, I_s = 23, V_s = 16, T_s = 41$; then the calibration factor, c_s, and inverse malfunction factor k_s are calculated as follows:

$$c_s = \frac{P_s \cdot V_s \cdot I_s}{T_s \cdot M_s} = 0.3902439,$$

and

$$k_s = P_s/M_s = 1/23 = 0.0434783.$$

The first subregion is chosen such that its center will be located at the voxel exhibiting the maximum net malfunction factor, as shown in Fig 5.2.

For the subregion, the number of voxels and the number of tests will change and they are calculated as follows, $V_1 = 7, T_1 = 29$. Since the inverse malfunction constant is a fixed value, i.e., $k_s = k_1$, the probability of lesion for the new subregion

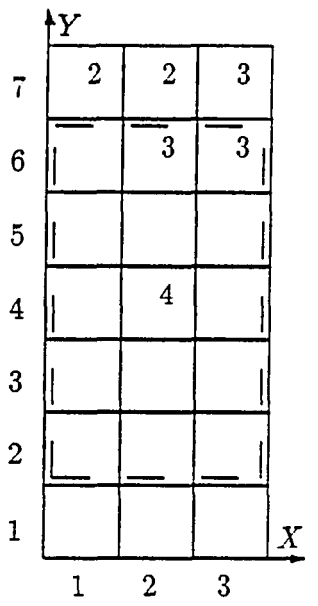


Figure 5.2: Sub–region R_1 of region R_s of Fig. 5.1. Enclosed with dashed lines is a sub–region R_2, also displayed in Fig. 5.3

can be calculated as follows:

$$P_1 = k_1 \cdot M_1 = k_s \cdot M_1 = \frac{1}{23}17 = 0.7391304.$$

As seen above, the total net malfunction factor for this subregion is calculated as $M_1 = 17$.

The calibration factor, c_1 for the subregion is calculated as

$$c_1 = \frac{k_s \cdot V_1 \cdot I_1}{T_1}$$

i.e.,

$$c_1 = \frac{(1/23) \cdot 7 \cdot 17}{29} = 0.1784108$$

The dimension of the new subregion will be half of the previous subregion and, again, the voxel with the maximum net malfunction factor is chosen as the center of the new subregion. The Fig. 5.3 will show this new subregion.

With the same approach as before, the probability of lesion and the calibration

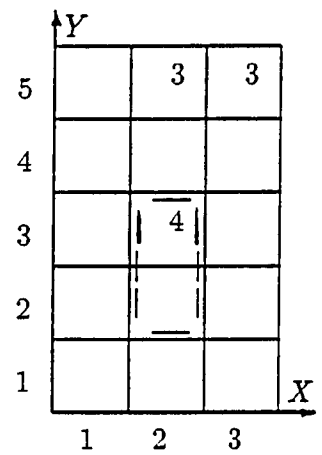


Figure 5.3: Sub–region R_2 of sub–region R_1 of Fig. 5.2. Enclosed with dashed lines is a sub–region R_3, which in turn, conatins a voxel (Region R_4) with $M = 4$.

factor for this smaller subregion are found to be

$$P_2 = k_s \cdot M_2 = \frac{1}{23}10 = 0.4347826$$

and

$$c_2 = \frac{k_s \cdot V_2 \cdot I_2}{T_2} = 0.1242236$$

where $V_2 = 4, T_2 = 14$ and $I_2 = 10$.

A third subregion contains a total of 2 voxels and 11 tests. Therefore, the calculation is

$$P_3 = k_s \cdot M_3 = \frac{1}{23}4 = 0.173913$$

and

$$c_3 = \frac{\frac{1}{23} \cdot 2 \cdot 4}{11} = 0.0316206$$

Finally, the voxel with the maximum net malfunction factor is defined as the new subregion, with $V_4 = 1, T_4 = 8$ and $M_4 = 4$. The following values are obtained for P_4 and c_4,

$$P_4 = \frac{1}{23}4 = 0.173913,$$

and

$$c_4 = \frac{\frac{1}{23} \cdot 1 \cdot 4}{8} = 0.0217391$$

Table 5.1 displays all the slope and calibration factors determined for the sequence of subregions outlined in the preceding discussion.

P_n	M_n	k_n	c_n
$P_0 = 1.0$	$M_0 = 23$	$k_0 = 1/23$	$c_0 = 0.3902439$
$P_1 = 0.7391304$	$M_1 = 17$	$k_1 = 1/23$	$c_1 = 0.1784108$
$P_2 = 0.4347826$	$M_2 = 10$	$k_2 = 1/23$	$c_2 = 0.1242236$
$P_3 = 0.173913$	$M_3 = 4$	$k_3 = 1/23$	$c_3 = 0.0316206$
$P_4 = 0.173913$	$M_4 = 4$	$k_4 = 1/23$	$c_4 = 0.0217391$

Table 5.1: Lesion Probabilities and Calibration Factors for region and sub–regions of Figures 5.1–3.

5.2 The Mean as the "Center of Lesion"

The **mean** of a data set is defined as the sum of the data elements divided by the number of pieces of data.

In the following example the mean and the variance of the whole region is calculated by using the relative malfunction frequency in the region.

The region used is shown in Fig. 5.1. In this case, the total positive net malfunction factor is, $M_s = 23$, as it was calculated before.

First the mean of the x and y distances are calculated as follows;

$$\mu_x = \left[2 \cdot 1 + 3 \cdot 1 - 5 \cdot 2 + 2 \cdot 2 + \cdots + 3 \cdot 4\right] \cdot \frac{1}{23} = 3.4782609 \approx 3,$$

$$\mu_y = \left[8 \cdot 2 + 9 \cdot 1 + 5 \cdot 4 + 7 \cdot 3 + \cdots + 9 \cdot 2\right] \cdot \frac{1}{23} = 7.3913044 \approx 7,$$

To find the variance, the following definition is used,

$$\mathrm{Var}(x) = \sum_{i=1}^{n} (\vec{X} - \mu_x)^2 f(x) \text{ and } \mathrm{Var}(y) = \sum_{i=1}^{n} (\vec{Y} - \mu_y)^2 f(y)$$

where μ_x and μ_y are the means of the locations X and Y as calculated above, and $f(x) = f(y) = M_v/M_s$ is the relative frequency of occurrence for each voxel involved.

Defining;

RMF = Relative Malfunction Frequency = $M_v/M_s = f(x) = f(y)$

$y =$ the Y distance of a voxel from the origin

$\mu_y =$ Mean of the distance in the y-direction

$\mathbf{V}(y)$ = Variance of the distance in the y-direction.

Table 5.2 shows the results of the calculations.

From Table 5.2, the Variances of y and x are calculated as $\mathbf{Var}(y) = 37.478262/23 = 1.6294896$, $\mathbf{Var}(x) = 19.739131/23 = 0.8582231$, and the standard deviations

Voxel $\vec{X}$	$\vec{Y}$	RMF	$(\vec{X}-\mu_x)$	$(\vec{Y}-\mu_y)$	$(\vec{X}-\mu_x)^2 RMF$	$(\vec{Y}-\mu_y)^2 RMF$
2	8	2/23	-1.4782609	0.6086957	4.3705104	0.0322183
2	9	1/23	-1.4782609	1.6086957	2.1852552	0.1125175
3	5	4/23	-0.4782609	-2.3913044	0.9149338	0.9944933
3	7	3/23	-0.4782609	-0.3913044	0.6862005	0.0199721
3	8	2/23	-0.4782609	0.6086957	0.4574467	0.0322183
3	9	1/23	-0.4782609	1.6086957	0.2287335	0.1125175
4	7	3/23	0.5217391	-0.3913044	0.8166352	0.0199721
4	8	3/23	0.5217391	0.6086957	0.8166352	0.0483275
5	8	2/23	1.5217391	0.6086957	4.63133799	0.0322183
5	9	52/23	1.5217391	1.6086957	4.63133799	0.2250349

Table 5.2: Computation of Variance for Example of Sub–Regions of Figure 5.1.

will be, $s.d_y = 1.2765146$ and $s.d_x = 0.9264033$ respectively.

The center voxel is chosen with $\vec{X} = 3$ and $\vec{Y} = 7$ and the similar calculations are done as before. Figure 5.4 shows the subregion, the voxels involved in the region, and the table displaying the results of the calculations done for this subregion. In this case, the means of y and x are found to be $\mu_{y_1} = 2.8125$ and $\mu_{x_1} = 2.125$.

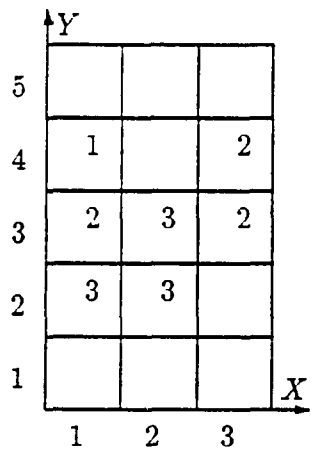


Figure 5.4: The sub–region of region shown in Fig. 5.1.

The variances are $\mathbf{Var}(y) = 0.5273438$, $\mathbf{Var}(x) = 0.671875$ and the standard

Voxel $\vec{X}$	$\vec{Y}$	RMF	$(\vec{X}-\mu_x)$	$(\vec{Y}-\mu_y)$	$(\vec{X}-\mu_x)^2 RMF$	$(\vec{Y}-\mu_y)^2 RMF$
1	2	3/16	-1.125	-0.8125	0.2373047	0.1237793
2	3	2/16	-0.125	0.1875	0.0029297	0.0043945
1	4	1/16	-1.125	1.1875	0.1582031	0.0881348
2	2	3/16	0.125	-0.8125	0.0029297	0.1237793
3	3	3/16	0.875	0.1875	0.0957031	0.0065918
1	3	2/16	-1.125	0.1875	0.0791016	0.0043945
3	4	2/16	0.875	1.1875	0.0957031	0.1762695

Table 5.3: Computation of Variance for sub–region of Fig. 5.4.

deviations are $s.d_y = 0.7261844$ and $s.d_x = 0.8196798$. For the subregion the sum of the positive net malfunction factors is $M_1 = 16$. Thus, we converge to the center of lesion at $\vec{X} = 2$ and $\vec{Y} = 2$ with $M = 3$.

5.3 Discussion on the median of a linear probability density model

The **median** of a data set is defined as the data value exactly in the middle of its ordered list if the number of pieces of data is odd, or it is the average of the two middle items if the number of pieces of data is even. The region used in this discussion is shown in Fig. 5.1.

To find the median, the $\vec{X}$ and $\vec{Y}$ distances are ranked in an increasing order with their respective relative malfunction frequencies.

$\vec{X}$	2_8	2_9	3_5	3_7	3_8	3_9	4_7	4_8	5_8	5_9
$\vec{Y}$	8_2	9_2	5_3	7_3	8_3	9_3	7_4	8_4	8_5	9_5
RMF	2/23	1/23	4/23	3/23	3/23	1/23	3/23	3/23	2/23	2/23

where 2_8 is $\vec{X} = 2$ and $\vec{Y} = 8$, that is the location of a voxel. From the frequency distribution above, the median of X will be $\vec{X} = 3$ and the median for Y will be $\vec{Y} = 8$. This voxel would be used as the center for any subsequent subregions to be chosen in the upcoming steps. In the following tables, the results of the calculations will be summarized, which is the appropriate compactification of previous RMF distribution.

Voxel $\vec{X}$	$\vec{Y}$	RMF	$(\vec{X} - \hat{X})$	$(\vec{Y} - \hat{Y})$	$\lvert(\vec{X} - \hat{X})\rvert f(x)$	$\lvert(\vec{Y} - \hat{Y})\rvert f(y)$
2	5	3/23	-1	-3	0.1304348	0.5217391
3	7	10/23	0	-1	0	0.2608696
4	8	6/23	1	0	0.2608696	0
5	9	4/23	2	1	0.3478261	0.173913

Table 5.4: Computation of the average deviations for region of Fig. 5.1.

The sum of the entries in the last columns are the *average deviations* for x and y, which are $17/23 = 0.739$ and $22/23 = 0.9565$, respectively. What this means is that, the boundaries of the new subregion can be defined with its center located at $\vec{X} = 2$ and $\vec{Y} = 2$, which is shown below.

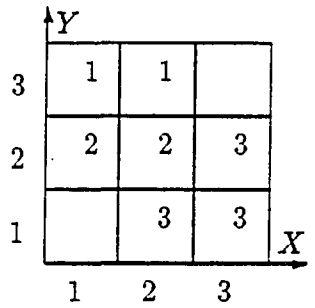


Figure 5.5: The subregion R_1 with Net Malfunction Factors

With the same approach as before, the $\vec{X}$ and $\vec{Y}$ distances are ranked as

$\vec{X}$	1_2	1_3	2_1	2_2	2_3	3_1	3_2	$med(x) = 2$
$\vec{Y}$	2_1	3_1	1_2	2_2	3_2	1_3	2_3	$med(y) = 2$
RMF	2/15	1/15	3/15	2/15	1/15	3/15	3/15	

which gives the median of $\vec{X}$ as 2, and the median of $\vec{Y}$ as 2. Thus, we converge to the center of lesion: $(\vec{X}, \vec{Y}) = (2, 2)$ for which the net–malfunction factor is $M(2, 2) = 2$.

5.4 Conclusions

This chapter focuses upon a variety of centralization techniques in order to "zero in" on regions with high probabilities of lesion. This hints at the possible existence of what will be referred to as a "center of lesion", a concept here explored from a simulation perspective which will be carefully pursued from a deterministic standpoint in Chapter 6.

It is known that:

a) if x_m is the mode of the data set $\{x_1, \cdots, x_N\}$ then $\hat{x} = x_m$ minimizes

$$\sum_{i=1}^{N} \Big[freq(x_i) - freq(\hat{x}) \Big]$$

b) if $\bar{x}_m$ is the arithmetic mean of the data set $\{x_1, \cdots, x_N\}$ then $\hat{x} = \bar{x}_m$ mini-

mizes

$$\sqrt{\sum_{i=1}^{N}\left[x_i - \hat{x}\right]^2 f_i}$$

c) if $\tilde{x}_m$ is the median of the data set $\{x_1, \cdots, x_N\}$ then $\hat{x} = \tilde{x}_m$ minimizes

$$\sum_{i=1}^{N}\left|x_i - \hat{x}\right| f_j$$

For these reasons, one selects a measure of central tendency appropriate to the needs of the problem. We have considered all of these three measures in our discussion and chose to use the arithmetic mean in the simulation because a major concern for us was to weigh all large Net–Malfunction Factors heavily in the process of subregion determination.

Chapter 6

A LOGISTIC MODEL

6.1 Introduction

We now pursue a deterministic approach to the problem of lesion localization. This deterministic modelling may be enhanced into a stochastic analysis when coupled with a Monte Carlo Simulation.

We intuitively observe that the distribution of lesion probability as a function of the overall sum of the positive net malfunction factors (M) must possess certain trends: both asymptotic and relating to the curvature of the graph. We intuitively recognize that as $M \to 0$ the distribution of lesion probability approaches 0 and as $M \to \infty$ the distribution of lesion probability approaches to 1 (for the universe).

How this distribution progresses as M increases from 0 to ∞ is also important. We believe that it must be monotone increasing. We are also concerned with its rate of increase. If it increases gradually then the underlying disease associated with this probability of lesion distribution does not have critical values of M. If it increases dramatically, for some relatively small increase in M, then the underlying neurological disease associated with this particular probability of lesion distribution possesses the characteristic of exhibiting dramatic, critical changes at some point after its onset.

Whether or not this dramatic change in lesion probability exists is determined by the disease and patient attributes (genetic.environmental and symptomatic). This tendency for dramatic increase in probability of lesion represents a critical case for neurologists that warrants further study. Our approach will be algebraic, probabilistic and will make extensive use of the calculus. In essence, we need to explore a technique for associating patient and disease attributes to critical increase in probability of lesion. In order to accomplish our task we must:

(1) pose a model (discussed in Section 6.2);

(2) justify its appropriateness (discussed in Section 6.2);

(3) study how certain parameters in the model affect its graph (discussed in Section 6.4);

(4) associate these parameters with patient and disease attributes (discussed in Section 6.5 and 6.5.3);

(5) focus upon the case in which a critical jump in lesion probability occurs (discussed in Section 6.6);

(6) estimate the size of the lesion probability jump in terms of the patient and disease attribute parameters (discussed in Section 6.6.1);

(7) associate this overall jump in lesion probability with its causative criteria: i.e., locate an impulse (or impulses) of lesion which generate this critical jump in lesion probability. (discussed in Section 6.8).

We ask the reader to bear with us in the following mathematical development which will provide us with insight into the above concern.

6.2 Justification For Model Selection

The model chosen is given as:

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

which can be shown graphically as follows,

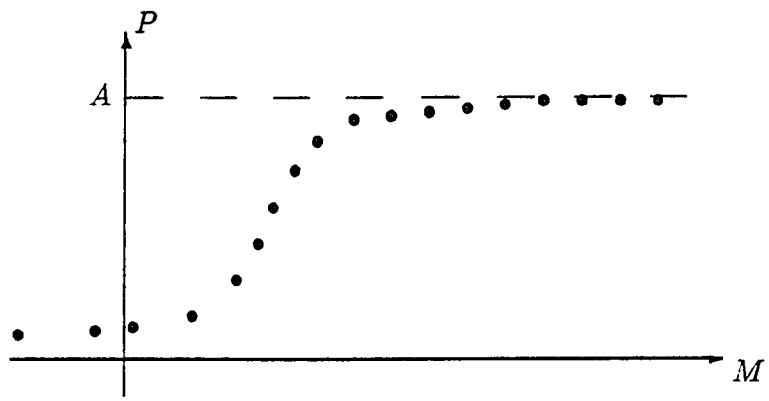


Figure 6.1: Logistic Sigmoid Model

This model is plausible because it contains intrinsic and necessary patterns in the rate of change of P with respect to M (i.e., the slope of the tangent to the curve).

These patterns are clear by inspection of the curve's convexity. Since one may observe that, for $0 \leq M < \frac{lnB}{\alpha}$, one has $P'' > 0$ indicating that the curve is concave up, it is evident that P' is an increasing function (i.e., the tangent to the curve is becoming increasingly more vertical as M moves to the right). There is a change when $M = \frac{lnB}{\alpha}$. This point is an *inflection point* because P'' changes sign (from positive values when $M < \frac{lnB}{\alpha}$ to negative values when $M > \frac{lnB}{\alpha}$). Thus, for

$$\frac{lnB}{\alpha} < M \leq M_{\text{Max. of the Universe}} < \infty$$

it is clear that since P'' is negative, and hence, P' is a decreasing function (i.e., the tangent to the curve is approaching a horizontal line). It is clear that, despite P being a monotone increasing function of M (with P' always positive for $M \geq 0$), the asymptotic trends of the curve agree with appropriate expectations, since P increases at the different rates which are demonstrated above.

Observe that $P = A(1 - Be^{-\alpha M})^{-1}$ is a solution to the following Initial Value Problem:

Initial Value Problem
Derivative Condition $$\frac{dP}{dM} = P' = \alpha A\left(\frac{P}{A}\right)\left(1 - \frac{P}{A}\right) \quad (6.1)$$
Initial Value $M = 0 \Rightarrow P = \frac{A}{1 - B}$ where $P = P(M)$ for $M \geq 0$.

Let us study this rate of change condition in Eq. 6.1.

As $P \to 0, P' \to 0$. Also, as $P \to A, P' \to 0$. Thus, the curve approaches horizontal asymptotes as $P \to 0$ and $P \to A$, since it approaches these values but never assumes them.

Also observe that P and P' are well defined for all values of M. The function P is bounded, and in fact the derivative is bounded since, first of all,

$$\begin{matrix} \alpha > 0 \\ A > 0 \end{matrix} \text{ and } 0 \leq P \leq A \Rightarrow P' \geq 0$$

and further, by deriving all possible inflection points (as shown below), the slope of the tangent,P', achieves its maximum value for $P = A/2$ (or $M = (lnB)/\alpha$). By computation, we see this value to be,

$$P'_{MAX} = \frac{\alpha A}{4} \text{ for } P = \frac{A}{2}$$

Thus

$$0 < P' \leq \frac{\alpha A}{4}$$

Figure 6.2: The Asymptotes of the Sigmoid Function

6.3 Points of Inflection: A Discussion

If

$$P' = \alpha A\left(\frac{P}{A}\right)\left(1 - \frac{P}{A}\right)$$

then

$$P'' = \alpha A\left[-\frac{PP'}{A^2} + \frac{P'}{A} - \frac{PP'}{A^2}\right]$$

and for $P'' = 0$ one gets,

$$\frac{P'}{A} = \frac{2PP'}{A^2}$$

$$A \cdot P' \cdot [A - 2P] = 0$$

Thus. the possible inflection points will be listed as follows;

$A = 0$	$P' = 0$	$A/2 = P$
Trivial case $P = 0$ always	This occurs for $P = 0$ or A. (see below)*	$P = A/2$ for $M = lnB/\alpha$

*(The curve never reaches these values because it approaches them asymptotically).

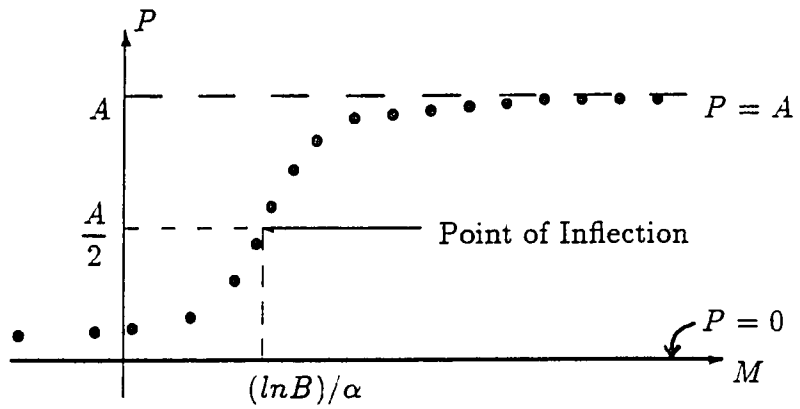


Figure 6.3: Points of Inflection of the Sigmoid Function

6.4 Variations in Parametric Values

Given.

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

with

$$P' = \alpha A\left(\frac{P}{A}\right)\left(1 - \frac{P}{A}\right)$$

and

$$P'' = \alpha P'\left[1 - \frac{P}{A/2}\right]$$

By inspection of the function and derivative relationships we see that.

(i) A is primarily a parameter that determines the spread of the plot.

Since one can write $P' = \alpha P(1 - P/A)$, we see that A primarily regulates when the asymptotic trend $P' \to 0$ is achieved.

(ii) In the way in which B is present in the equation for P, large values of B always cause the function $P = P(M)$ to have drastic convexity.

(iii) Clearly, α very large produces diminished convexity beyond all effects of B. However α very small produces a curve which approaches a linear trend (This

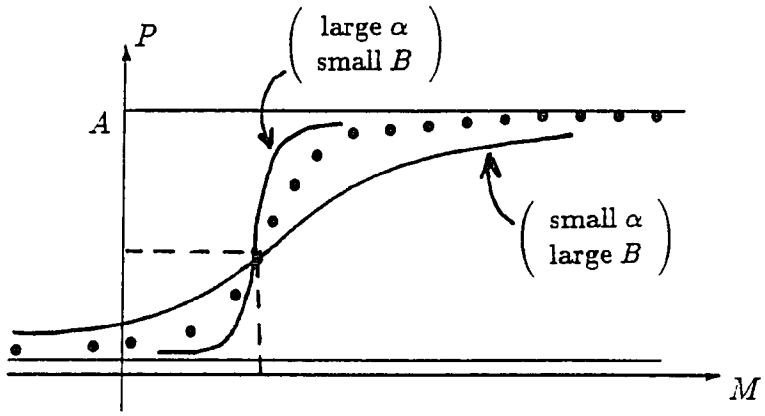


Figure 6.4: Asymptotic trends of the Sigmoid Function

is clear since

$$P'' = \alpha P'[1 - P/(A/2)]$$

is very small for α very small which means that the convexity is very small).

That is.

$$\begin{aligned} & P'' \approx 0 \\ \Rightarrow \quad & P' \approx \text{ constant} \\ \Rightarrow \quad & P \approx \text{ linear} \end{aligned}$$

Which values of A, B, α yield a sequence of models which reflect certain expectations of neurologists ? A general statement would be: The function is more realistic for α small, not too small; B large, not too large; with A serving as a maximum probability of lesion which defines the overall spread of the lesion probability distribution. From the trial runs attached it can be seen that $\alpha = 0.2$, $P_0 = 0.9$, $A = 1$, $B = 33.03$ produce a very plausible model.

6.5 Examples

In the following, the sigmoid logistic model has been analyzed for different values of the parameters involved in the model. The model is given as follows:

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

For this model, B is defined as;

$$B = c \cdot \frac{(T)}{(V)} \cdot \frac{1}{I}$$

where c is the *calibration factor* which is calculated from the á priori probability conditions of the section under study; T is the total number of tests applied in the section; and V is the total number of voxels involved in the section. I is the involvement factor taken as $I = \sqrt{M^2 + F^2}$; and M is defined as the sum of positive net malfunction factors in the section under study. In the preliminary calculations, A, which is the maximum lesion probability, is set to unity.

In Table 6.1[1], we list, for different values of α and A, the subsequent lesion probabilities for appropriate subregions.

The attached graphs are in two sets. Each graph depicts the Probability of lesion versus the Positive Net Malfunction Factor. In one set of graphs, the maximum lesion probability is held constant and the parameter α is varied accordingly. In the other set, for a fixed parameter α, the maximum lesion probability is allowed to vary.

[1] The region used is the left half of Section 7, namely voxels 700, ..., 704; 710, ..., 714, 790, ..., 794

M	P $\alpha = 0.1$	P $\alpha = 0.2$	P $\alpha = 0.25$	P $\alpha = 0.5$	P $\alpha = 0.8$
0	0.0060765	0.0003364	7.8958E-05	5.6E-08	9.3353E-12
8	0.0134236	0.001664	5.8312E-04	3.0596E-06	5.6E-09
10	0.0163471	0.0024803	9.6103E-04	8.3167E-06	2.78E-08
15	0.0266689	0.0067137	3.3464E-03	1.0131E-04	1.5304E-06
20	0.0432219	0.018042	1.1583E-02	1.2328E-03	8.2949E-05
25	0.0693173	0.047567	3.9297E-02	1.4919E-02	4.5088E-03
29	0.1	0.1	0.1	0.1	0.1
0	0.0135692	0.00075633	1.7751E-04	1.2161E-07	2.1005E-11
8	0.0368572	0.0037349	1.3102E-03	6.8839E-06	1.26E-08
10	0.0446531	0.0055617	2.1583E-03	1.8712E-05	6.26E-08
15	0.0580693	0.014975	7.4928E-03	2.2791E-04	3.4434E-06
20	0.0922644	0.039685	2.5673E-02	2.7695E-03	1.8662E-04
25	0.1435277	0.10099	8.4224E-02	3.2726E-02	1.0088E-02
29	0.2	0.2	0.2	0.2	0.2
0	0.05211536	3.0195E-03	6.9589E-04	5.043E-07	8.4019E-11
8	0.1090968	1.4779E-02	5.2236E-03	2.7535E-05	5.06E-08
10	0.1301085	2.1889E-02	8.5831E-03	7.4845E-05	2.505E-07
15	0.1978161	5.7347E-02	2.8759E-02	9.1103E-04	1.3773E-05
20	0.2890505	0.14189	9.5407E-02	1.0987E-02	7.4605E-04
25	0.4013123	0.301	0.2651	0.11996	3.9166E-02
29	0.5	0.5	0.5	0.5	0.5
0	0.3311972	2.6535E-02	6.3553E-03	4.5263E-06	7.5614E-11
8	0.524287	0.11895	4.5127E-02	2.4704E-04	4.551E-07
10	0.5737639	0.16765	7.2285E-02	6.7132E-04	2.2541E-06
15	0.6893805	0.3538	0.2138	8.1173E-03	1.2394E-04
20	0.7853675	0.59811	0.48698	4.066E-02	5.3467E-03
25	0.8578108	0.8018	0.76815	0.55022	0.26839
29	0.9	0.9	0.9	0.9	0.9

Table 6.1: The list of lesion probabilities vs M for different values of α and A.

The values of Table 6.1 are plotted in Figures 6.5–6.9, they are graphs of

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

with M on the horizontal axis, and P on the vertical axis. Four curves are shown on each page; individual curves for A values of: 0.1, 0.2, 0.5 and 0.9. The plots on each page have a fixed α value.

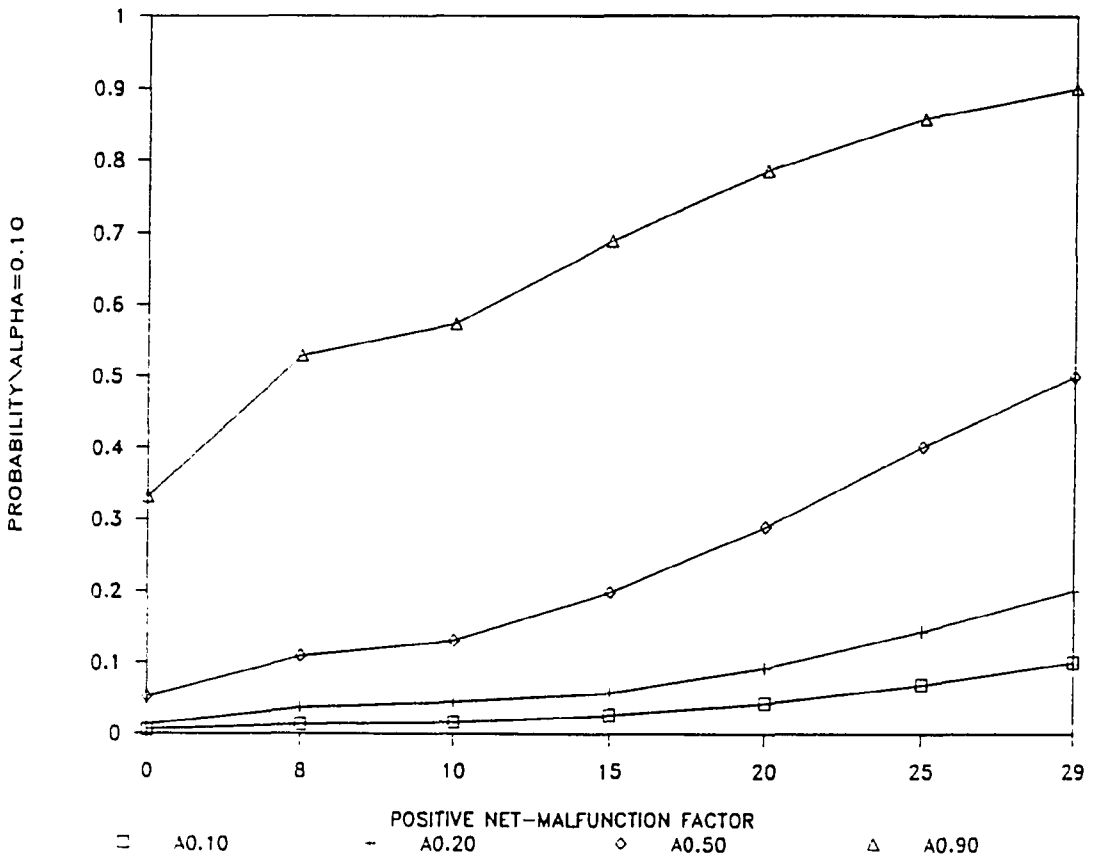


Figure 6.5: Probability vs Net-Malfunction Factor for $\alpha = 0.1$

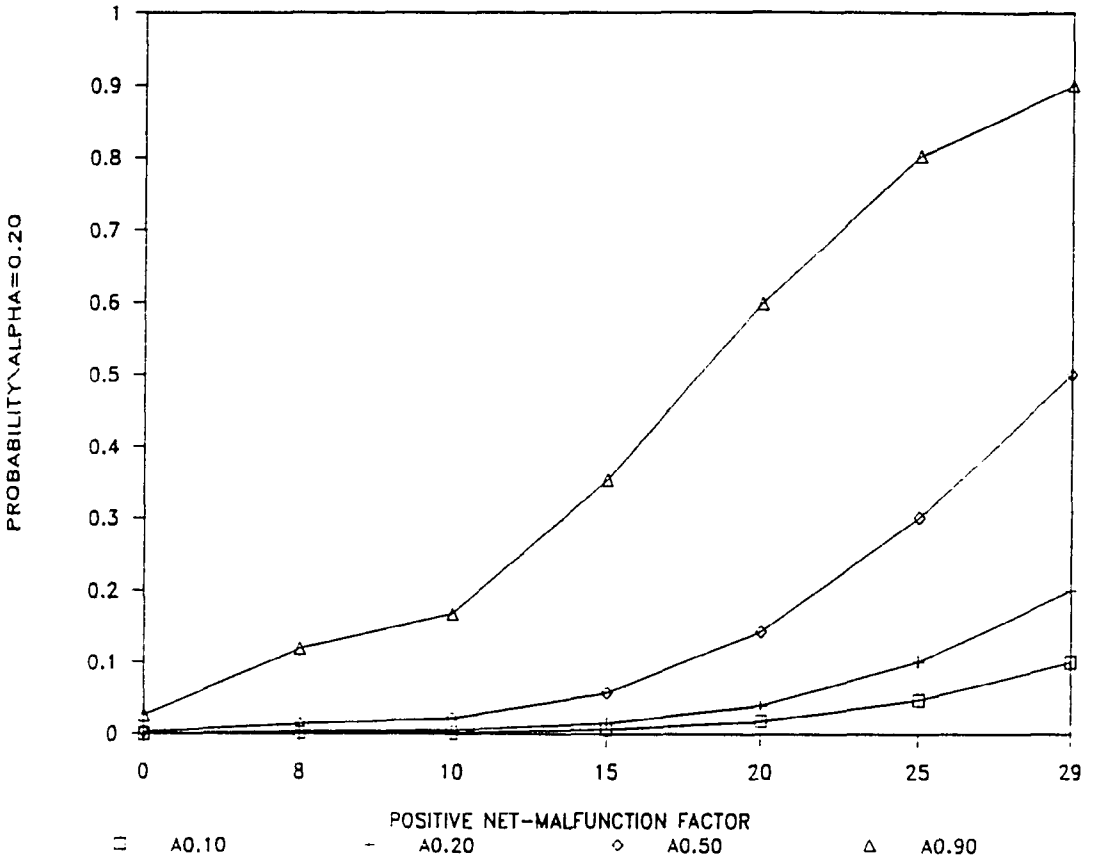


Figure 6.6: Probability vs Net-Malfunction Factor for $\alpha = 0.20$

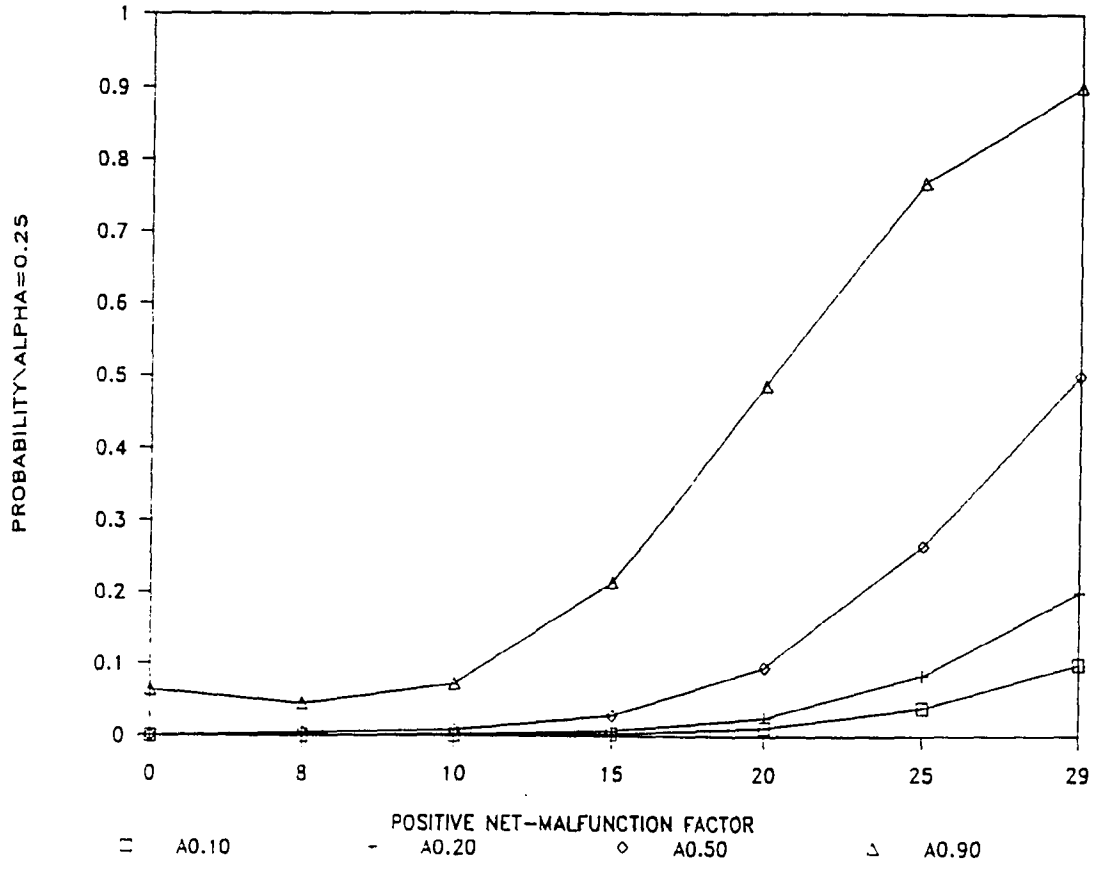


Figure 6.7: Probability vs Net-Malfunction Factor for $\alpha = 0.25$

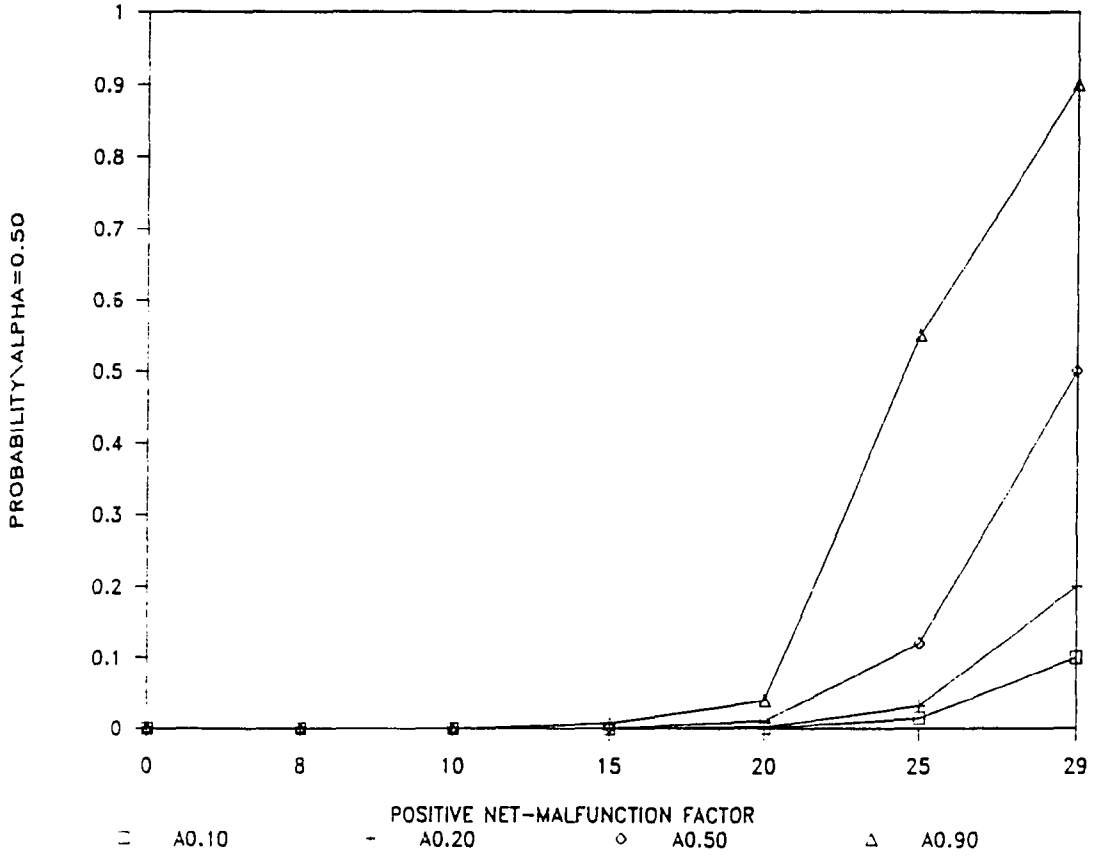


Figure 6.8: Probability vs Net-Malfunction Factor for $\alpha = 0.50$

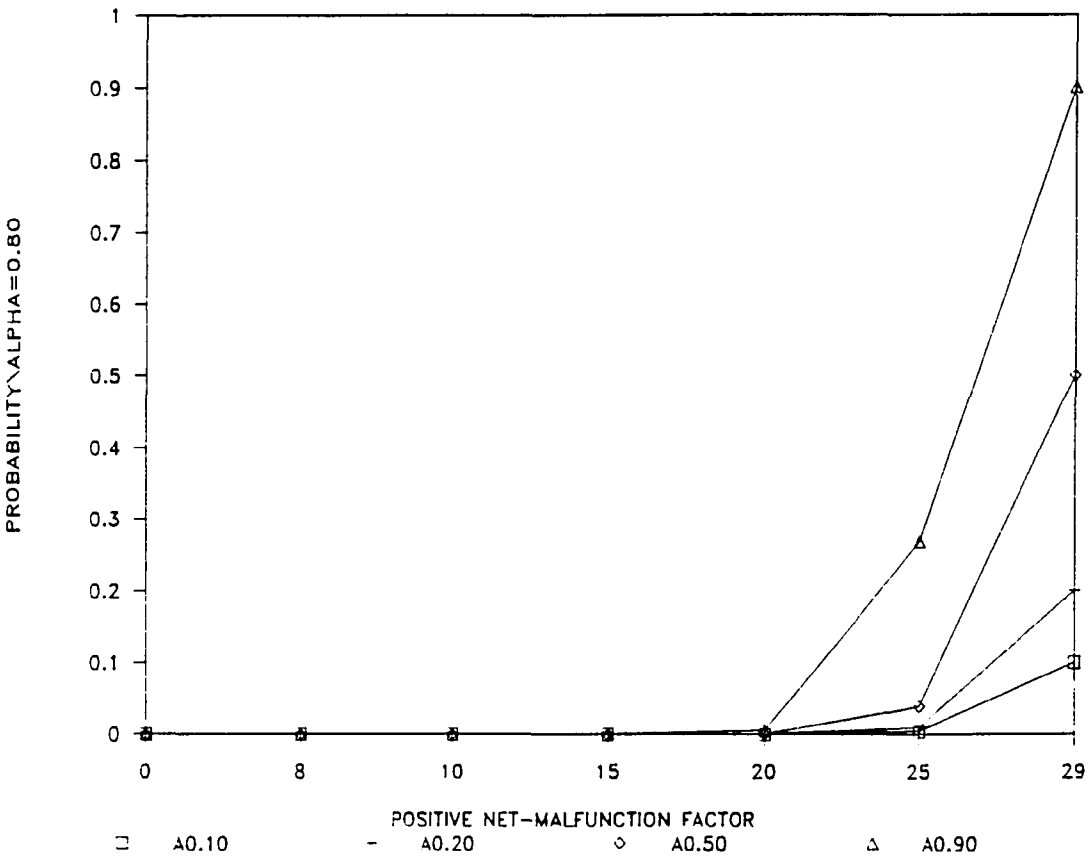


Figure 6.9: Probability vs Net-Malfunction Factor for $\alpha = 0.80$

6.5.1 Trends in the overall distribution of lesion probability effected by the parameter α

We observe that, as expected, large values of α (say, $\alpha = 0.80$) produce a family of curves that experience extreme convexity. There is evidence of a critical value for M which, when attained, causes the lesion probability to rapidly jump from a very small to a very large value. This jump is, of course, proportional to A. As α decreases, and reaches a value of 0.50, one begins to see traces of our desired schematic in an apparent change of convexity visible for large values of A.

Let us allow α to decrease further. We see that (for larger values of A) an α value of 0.20 produces a curve reflecting our expectations–since small changes in value of M should not produce drastically different lesion probabilities. Even smaller values of α produce a family of curves that approach linearity with a jump near $M = 0$.

The values of Table 6.1 are plotted in Figures 6.10–6.13, they are graphs of

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

with M on the horizontal axis and P on the vertical axis. Five curves are shown on each page; individual curves for α values of 0.1 0.2, 0.25, 0.50 and 0.80. The plots on each page have a fixed A value.

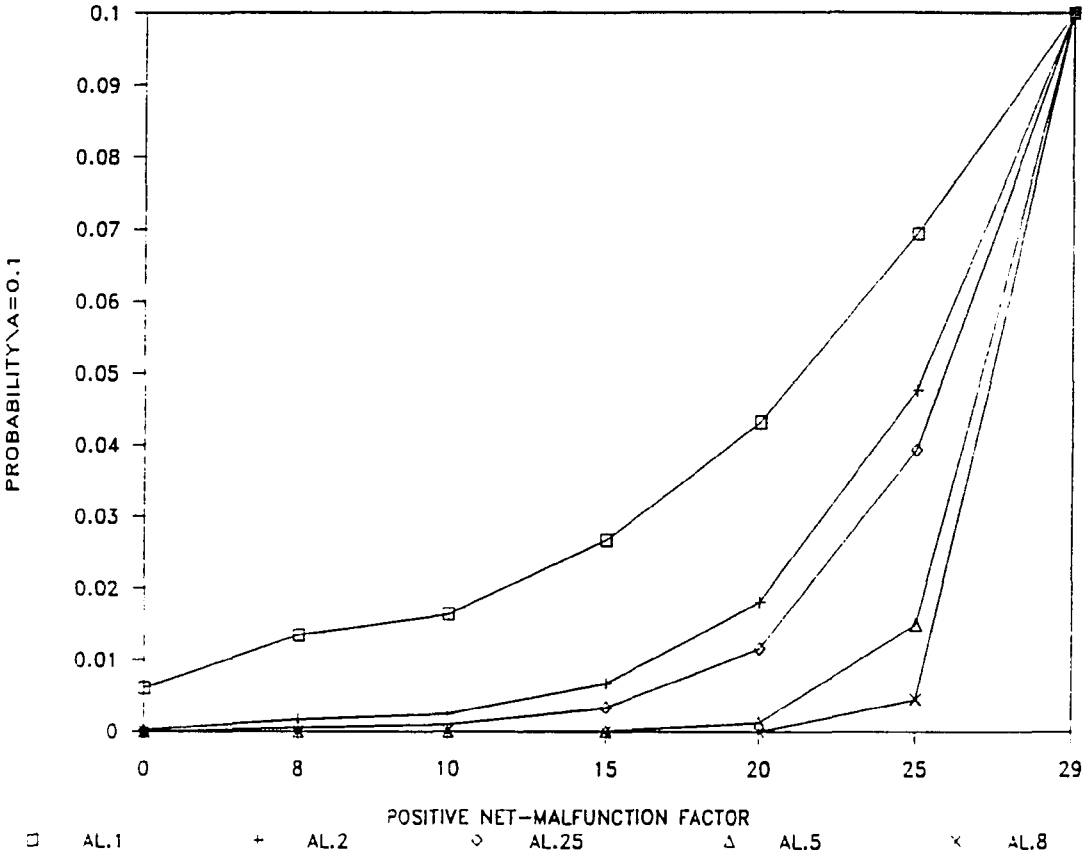


Figure 6.10: Probability vs Net-Malfunction Factor for $A = 0.10$

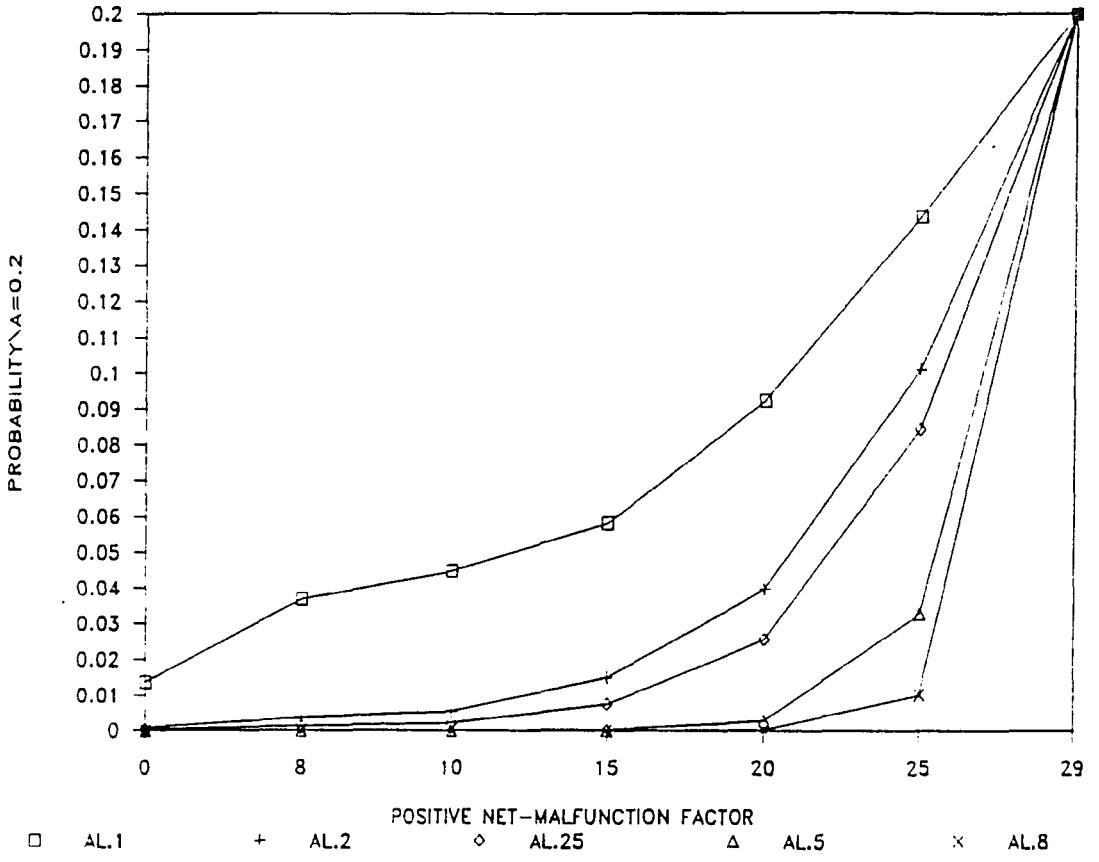


Figure 6.11: Probability vs Net-Malfunction Factor for $A = 0.20$

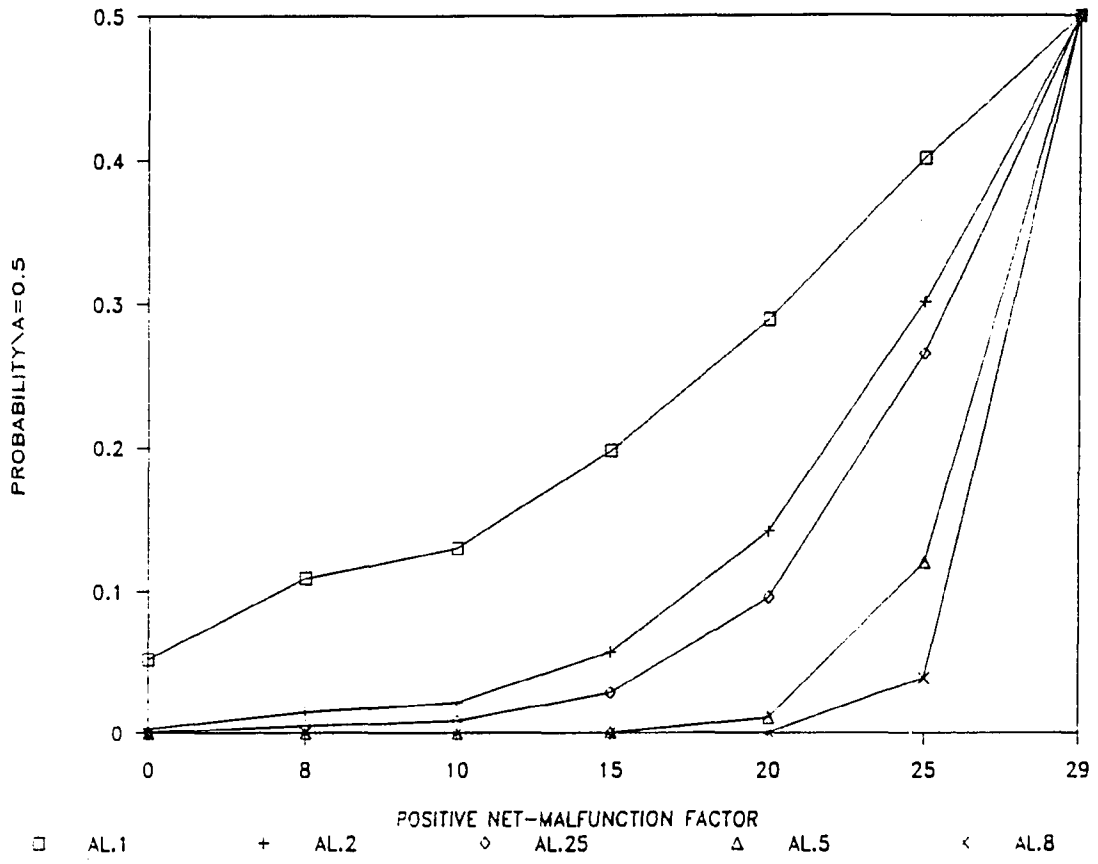


Figure 6.12: Probability vs Net-Malfunction Factor for $A = 0.50$

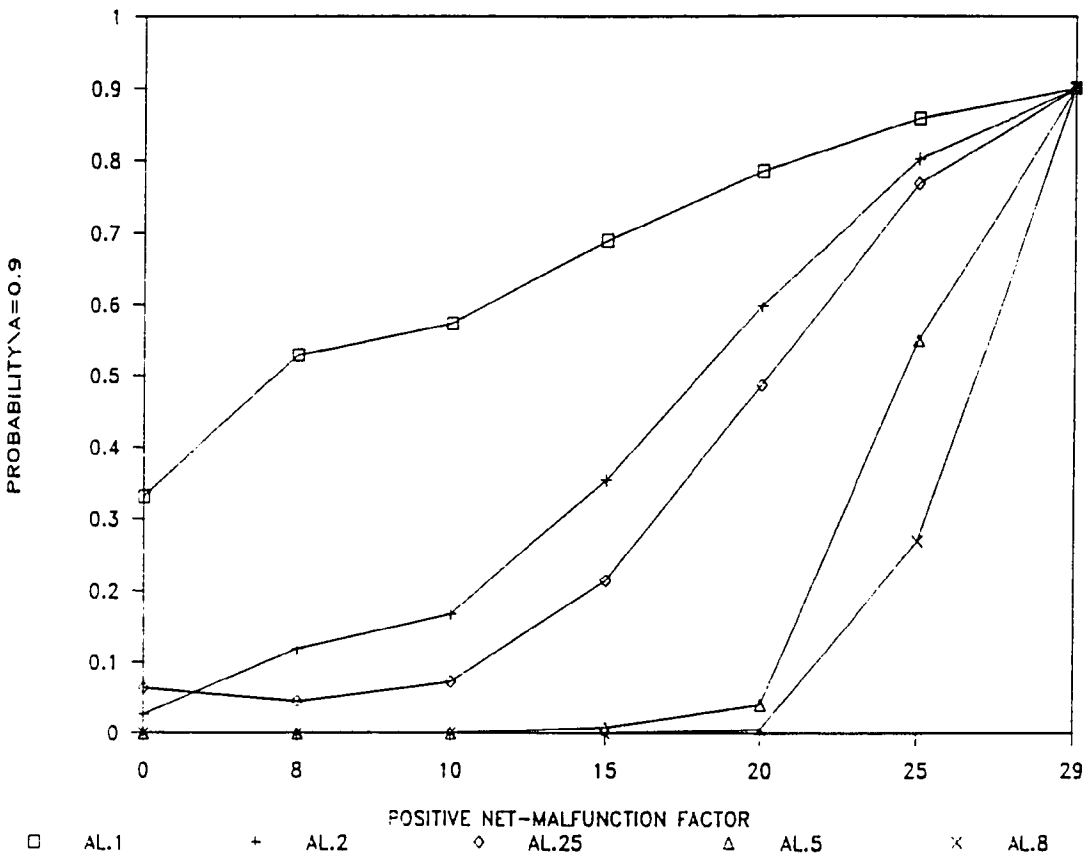


Figure 6.13: Probability vs Net-Malfunction Factor for $A = 0.90$

6.5.2 Trends in the overall distribution of lesion probability effected by the parameter, A

The desirable convexity trends for

$$P = \frac{A}{1 + Be^{-\alpha M}}$$

are apparent for larger values of A,($A = 0.90$), and α in the range of $0.1 \leq \alpha \leq 0.5$ The family of curves generated for fixed A values where $A \leq 0.5$ do not reflect the convexity changes which we seek.

6.5.3 Significance of the results on the variation of parametric values

Our motivation in pursuing a comparison of the logistic model curves for different values of α and A is based upon some intuitive understanding of the graphical charateristics of the overall probability of lesion as a function of M (M being the aggregate of all positive net malfunction factors). Exaggerated convexity which is attained for <u>α large</u>, indicates the presence of a critical value for M at which a large jump in lesion probability occurs; this may be considered a threshold value which, when attained, leads to an almost certain presence of lesion. When convexity changes are almost absent (i.e., <u>α small</u>) the overall lesion probability is almost linear–there is no presence of a threshold value, which suggests we should turn to the concept of a malfunction angle θ, as discussed in Chapter 4, to locate the transition point for M (i.e., the M values for which one crosses from the "ambiguous" to the "malfunction" regions).

The more speculative case is for moderate convexity in which α is midrange (say, $0.2 \leq \alpha \leq 0.8$); we see that the inflection point represents the M^* value ($M^* = lnB/\alpha$) which, when attained, indicates that the overall lesion probability is increasing rapidly rather than slowly as before. This M^* is somewhat of a threshold and it should represent to the neurologist, making use of this model, that $M > M^*$

represents entry into a cautious region in which the overall lesion probability may rapidly escalate with future malfunction test outcomes.

Small values of A correspond to neurological examinations focusing on anatomical regions or diseases with a minimal overall lesion probability. No threshold is apparent since there is little range for probability fluctuations.

The severity of the nature and outcome of a neurological examination may be partially viewed in terms of the chosen value of A as shown below.

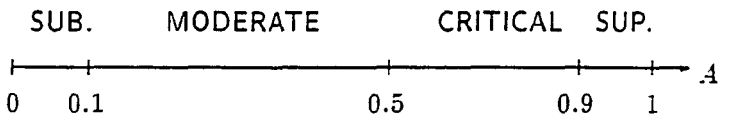


Figure 6.14: A values and corresponding level of severity

where. SUB=SUBCRITICAL. and SUP=SUPERCRITICAL respectively.

In actuality, future research may focus upon associating a set of values for the triad (α, B, A) with all acquired personal data for the patient undergoing the examination as well as the primary symptomalogy which is present (suggestive of a specific disease). This association process would involve extensive Monte Carlo simulation in conjunction which significant expansion of the neurological information database.

The physician would then be able to identify α, B and A for the patient he is about to further examine and actually see from the plot of $P(M)$ versus M, the presence or absence of threshold values presenting regions of lethality, as well as the overall limiting value of probability of lesion.

6.6 Step Function Approach

In this section, the sigmoid logistic model will be viewed as a step function generated by *lesion impulses*. It was observed that exaggerated convexity in the logistic model (i.e., α large) produces drastic change in the P'' polarity (from positive to negative). Recall that,

$$P'' = \alpha A\left[\frac{P'}{A} - \frac{2PP'}{A^2}\right]$$

and assume M is in the neighborhood of the inflection point: $P = A/2$ for $M = lnB/\alpha$. For $P \approx A/2$, the curve possesses exaggerated convexity in the following cases: at least one of the following are large: α, A, P'/A, $1 - 2P/A$; while the other remaining three quantities are not very small.

Consider the following figure,

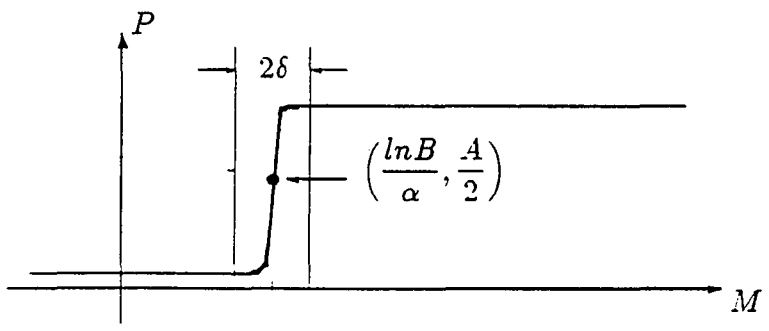


Figure 6.15: Exaggerated convexity of the Sigmoid Function

The original function is given as:

$$P = \frac{A}{1 + Be^{-\alpha M}}.$$

Here M_1 is in a δ neighborhood of $\frac{lnB}{\alpha}$,

$$\left|M_1 - \frac{lnB}{\alpha}\right| < \delta \text{ for } \delta > 0, \text{ small}$$

Suppose,

$$M_1 - \frac{lnB}{\alpha} = \epsilon_1 = \text{very small.}$$

or

$$M_1 = \epsilon_1 + \frac{lnB}{\alpha}$$

then

$$P_1 = \frac{A}{1 + Be^{-\alpha(\epsilon_1 + lnB/\alpha)}}$$

or

$$P_1 = \frac{A}{1 + Be^{-\alpha\epsilon_1}e^{-lnB}} = \frac{A}{1 + e^{-\alpha\epsilon_1}}$$

It is clear that

1 - α large allows $|P_1 - A/2|$ to be large, which is necessary for exaggerated convexity.

2 - A large also allows extreme convexity simply because $A/2$ and A are very far apart iff A is large.

3 - Consider P'/A. P'/A is large iff we are in a neighborhood of lnB/A (which is indeed the case!) and then it may be estimated as follows:

$$P' = \alpha A\left(\frac{P}{A}\right)\left(1 - \frac{P}{A}\right)$$

$$\frac{P'}{A} = \alpha\left(\frac{P}{A}\right)\left(1 - \frac{P}{A}\right)$$

$$\frac{P'}{A} \approx \alpha\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{\alpha}{4}$$

4 - $\left(1 - 2P/A\right)$ is small in a neighborhood of the inflection point and hence only very large values of α and/or A can produce a large $P''(M_1)$ if ε_1 is very small; but it is, never the less, possible to give appropriate α and A (lower bounds) for which any pre specified P'' value will be attained within an ε_1 neighborhood of the inflection point.

6.6.1 Estimation of change in lesion probability for large α

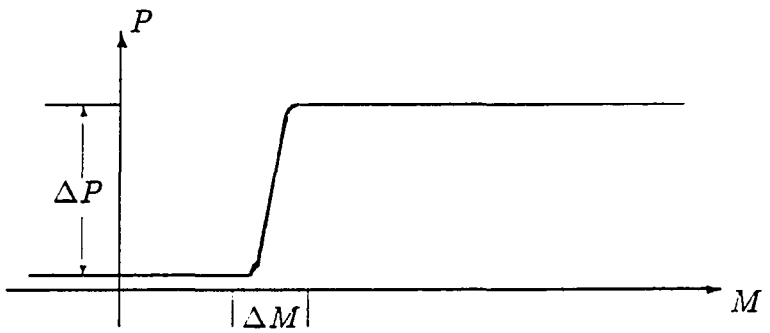


Figure 6.16: Changes in M and P

Let Q be the change in the probability,

$$\begin{aligned} \text{Change in vertical direction} &= Q = \Delta P \\ \text{Change in horizontal direction} &= 2\delta = \Delta M \end{aligned}$$

Q represents the change in the probability for some ΔM.

$$P_2 - P_1 = Q = \frac{A}{1 + Be^{-\alpha M_2}} - \frac{A}{1 + Be^{-\alpha M_1}}$$

But $M_2 = M_1 + 2\delta$.

$$P_2 - P_1 = Q = \frac{A}{1 + Be^{-\alpha(M_1+2\delta)}} - \frac{A}{1 + Be^{-\alpha M_1}}$$

After some simplifications

$$Q = \frac{ABe^{-\alpha M_1}\left[1 - e^{-2\alpha\delta}\right]}{(1 + Be^{-\alpha M_1})\left[1 + (Be^{-\alpha M_1})e^{-2\alpha\delta}\right]}$$

Let $\bar{x} = 1 + Be^{-\alpha M_1}$, then

$$Q = A\frac{(\bar{x} - 1)\cdot(1 - e^{-2\alpha\delta})}{\bar{x}\cdot(1 + (\bar{x} - 1)e^{-2\alpha\delta})}$$

Then, if we let

$$E = \frac{\bar{x}}{\bar{x} - 1} \quad \text{and} \quad F = \bar{x}$$

we have

$$Q = \frac{A(1 - e^{-2\alpha\delta})}{E + Fe^{-2\alpha\delta}}$$

It is clear that E and F are functions of α: but, they are independent of δ and, for the moment, our goal is to find an estimate of Q as a function of δ.

Allowing $e^{-2\alpha\delta}$ to be approximated by the first 3 terms of its Taylor Series ($e^{-2\alpha\delta} \approx 1 - 2\alpha\delta + 2\alpha^2\delta^2$) we find

$$Q \approx \frac{A(2\alpha\delta - 2\alpha^2\delta^2)}{E + F - 2\alpha\delta + 2\alpha^2\delta^2}$$

and for δ very small

$$Q = \frac{2A\alpha\delta}{E + F}$$

but, since

$$E = \frac{e^{\alpha M_1}}{B} - 1$$

and

$$F = 1 + Be^{-\alpha M_1}$$

we may estimate

$$E + F \quad \text{by} \quad \frac{e^{\alpha M_1}}{B} + 2$$

for α large.

Thus

$$Q \approx \frac{2A\alpha\delta}{\frac{e^{\alpha M_1}}{B} + 2} = \frac{2AB\alpha\delta}{e^{\alpha M_1} + 2B}$$

i.e., our estimate of Q is governed by the proportionality

$$Q \approx \left(\frac{2AB\alpha}{e^{\alpha M_1} + 2B}\right)\delta$$

Q represents the jump in lesion probability for a δ neighborhood between M_1 and M_2. If M_1 is chosen to be $M_1 = lnB/\alpha$, then

$$Q \approx \frac{2AB\alpha}{3B}\delta = \frac{2}{3}(A\alpha)\delta$$

is our estimate for the lesion probability increase.

6.7 Moments

6.7.1 Definitions

First consider the following definitions[31].

The Moment generating function:

$$M_X(\Theta) = E\left(e^{\Theta x}\right) = \int_{-\infty}^{\infty} e^{\Theta x} f(x)dx$$

The double sided Laplace transform:

$$F(s) = \mathcal{L}\{f(t)\} = \int_{-\infty}^{\infty} e^{-st} f(t)dt$$

The following properties of the Moment Generating function will be employed:

$$M_X(\Theta = 0) = 1$$

$$\frac{dM_X}{d\Theta}(\Theta = 0) = \mu$$

$$\frac{d^2 M_X}{d\Theta^2}(\Theta = 0) = \mu^2 + \sigma^2$$

[Please note that $M_X(\Theta)$ exits and is well defined if $f(x)$ is a probability density function: the Laplace transform of a probability density function is also well defined since these density functions are piecewise continuous and of exponential order (i.e., they can be bounded by an exponential with linear argument).]

6.7.2 Moment generating function for lesion probability density

Now recall our probability distribution

$$F(x) = \frac{A}{1 + Be^{-\alpha x}} = P\{X \leq x\}$$

as $\lim_{x \to \infty} F(x) = A = 1$ for the universe, and it is less than one for subregions of the universe. We observe that the density function which generates this distribution is

$$f(t) = \frac{AB\alpha e^{-\alpha t}}{(1 + Be^{-\alpha t})^2} \quad \text{for} \quad -\infty < t < \infty$$

where $A = 1$ for the universe. In addition, we require

$$\int_{-\infty}^{\infty} f(t)dt = 1$$

and indeed

$$\lim_{x \to \infty} \int_{-\infty}^{x} \frac{AB\alpha e^{-\alpha t}}{(1 + Be^{-\alpha t})^2} dt = \lim_{x \to \infty} \left(\frac{A}{1 + Be^{-\alpha x}} \right) = A = 1$$

Hence we may consider a double sided Laplace Transform,

$$\mathcal{L}\{f(t)\} = \int_{-\infty}^{\infty} e^{-st} f(t)dt$$

analogous to the Moment Generating function.

$$M_X(\Theta) = \int_{-\infty}^{\infty} e^{\Theta x} f(x)dx$$

Laplace Transforms exist for functions of exponential order which are piecewise continuous on the real line. Thus we see that:

$$\mathcal{L}\{f(t)\} = F(s) \quad \text{and} \quad M_X(f(x)) = M(\Theta)$$

so we associate

$$F(-s) = M(\Theta) \quad \text{or} \quad F(s) = M(-\Theta)$$

In order to generate $M(\Theta)$ for our lesion probability distribution we must compute

$$M_X(\Theta) = \int_{-\infty}^{\infty} e^{\Theta x} f(x)dx = \int_{-\infty}^{\infty} \frac{AB\alpha e^{(\Theta - \alpha)x}}{(1 + Be^{-\alpha x})^2} dx$$

and this result may be used to find μ and σ^2 for our lesion probability distribution.

6.7.3 Motivation

The distribution for lesion probability, $P(M) = A/(1 + Be^{-\alpha M})$, may be viewed as the moment generating function of a lesion impulse density for α large, (i.e., $P(M)$ is almost a step function).

Consider the concept of "center of lesion" equivalent to an impulse. When the lesion probability distribution function comes in contact with this location (center of lesion–identified by M^* value) the probability jumps from (ε_1) to $(1 - \varepsilon_2)$ within a 2δ interval.

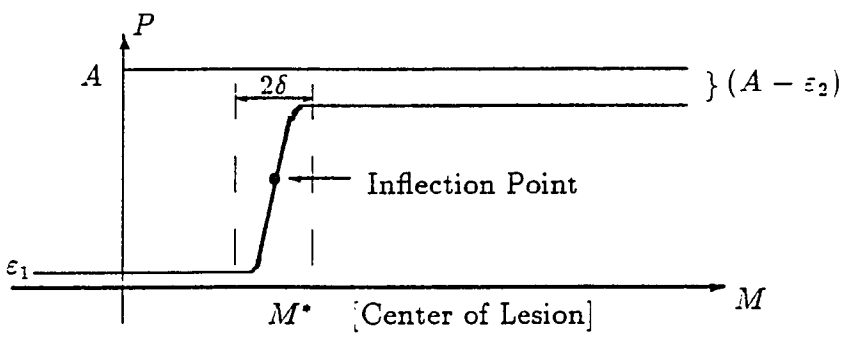


Figure 6.17: Center of Lesion

A Laplace transform operates on an impulse function to generate a step function as its end product. The following relation will be explored using the Laplace Transform properties: if a probability of lesion distribution can be modeled by a step function, then the inverse Laplace Transform of this distribution, if it exists, may approach an impulse (or sequence of impulses) which may be considered causative (from a neurologist's perspective). Hence, one may infer the presence of a lesion at such a location, and one may refer to the center of the impulse region as a center of lesion. Thus,

$$\mathcal{L}\{\text{Impulse function}\} = \mathcal{L}\{\mathcal{L}^{-1}\{F(M)\}\} = F(M) = \text{Step Function}$$

where

$$F(M) = \frac{A}{1 + Be^{-\alpha M}} = \mathcal{L}\{\text{Impulse Function}\}$$

Hence

$$\text{Impulse Function} = \mathcal{L}^{-1}\left\{\frac{A}{1 + Be^{-\alpha M}}\right\}$$

The explicit solution of this Inverse Laplace Transform is shown in section 6.7.

6.7.4 Observation

Given X, a random variable, SS_X, its corresponding sample space, and $f(x)$, its corresponding continuous probability density function, then the moment generating function is found by evaluating

$$M_X(\Theta) = \int_{SS_X} e^{\Theta x} f(x)dx$$

which becomes

$$M_X(\Theta) = \int_0^\infty e^{\Theta x} f(x)dx + \int_{-\infty}^0 e^{\Theta x} f(x)dx$$

if $SS_X = (-\infty, \infty)$. If we then let $u = -x$ we have,

$$M_X(\Theta) = \int_{-\infty}^0 e^{-\Theta u} f(-u)du + \int_0^\infty e^{-\Theta u} f(-u)du = \int_{-\infty}^\infty e^{-\Theta u} f(-u)du$$

Thus if we viewed this final expression as depicting some moment generating function we could conceptualize $-u = X$ as a random variable and $SS_{-u} = (-\infty, \infty)$ as the new sample space and

$$M_{-u}(\Theta) = \int_{-\infty}^\infty e^{-\Theta u} f(-u)du$$

or

$$M_{-u}(\Theta) = \int_{-\infty}^0 e^{-\Theta u} f(-u)du + \int_0^\infty e^{-\Theta u} f(-u)du$$

Now if we let $f(-u) = g(u)$ then,

$$M_{-u}(\Theta) = \mathcal{L}\{g(u)\}$$

Furthermore, if $g(u)$ is the density function for the lesion impulses, which are depicted as generating our step function, then

$$\mathcal{L}\{g\} = \frac{A}{1 + Be^{-\alpha M}} = M_{-u}(\Theta)$$

And thus we may compute μ and σ^2 for the impulse density function using the properties in section 6.6.1. That is,

$$\frac{d}{ds}\mathcal{L}\{g\} = \frac{d}{ds}P(s) = \frac{d}{ds}\left[\frac{A}{1 + Be^{-\alpha s}}\right]\Bigg|_{s=0}$$

$$\frac{d}{ds}P(s)\Big|_{s=0} = \frac{AB\alpha}{(1 + B)^2}$$

and since

$$\frac{d\mathcal{M}_{-u}}{d\Theta}\Big|_{\Theta=0} = \mu$$

then,

$$\mu = \frac{AB\alpha}{(1 + B)^2}$$

is the *centroid* for the lesion Impulse density function.

The variance will be calculated as follows:

$$\frac{d^2P(s)}{ds^2}\Big|_{s=0} = \mu^2 + \sigma^2$$

We find

$$\mu^2 + \sigma^2 = \frac{AB\alpha^2(B - 1)}{(1 + B)^3}$$

and hence eliminating μ^2

$$\sigma^2 = \frac{AB\alpha^2(B^2 - AB - 1)}{(1 + B)^4}$$

6.7.5 Series Expansion of $\mathcal{L}\{g(u)\}$

We wish to find the density for the lesion impulse(s). This is facilitated by a power series expansion of $P(M)$.

The model is given as follows;

$$P = \frac{A}{1 + Be^{-\alpha M}} \tag{6.2}$$

Recalling the geometric series theorem:

$$P(M) = \lim_{N \to \infty} \sum_{k=0}^{N} A(-1)^k \left[Be^{-\alpha M} \right]^k \tag{6.3}$$

provided that the following convergence condition holds:

$$\left| Be^{-\alpha M} \right| < 1$$

Let us explore this convergence condition. If P_0 is a value of $P(M)$ for which we have convergence then we require;

$$P_0 = \frac{A}{1 + B_0 e^{-\alpha M_0}} \tag{6.4}$$

where

$$B_0 = \frac{T_0 \cdot c_0}{V_0 \cdot I_0} \tag{6.5}$$

If we set $A = 1$ (as we do for the sample space)

$$P_0 \left[1 + B_0 e^{-\alpha M_0} \right] = A = 1$$

hence

$$B_0 e^{-\alpha M_0} = \frac{1 - P_0}{P_0},$$

and since we require $|B_0 e^{-\alpha M_0}| < 1$, we then find

$$\left| \frac{(1 - P_0)}{P_0} \right| < 1 \Rightarrow \frac{1}{2} < P_0 < 1$$

or, in general, $\frac{A}{2} < P_0 < A$ and, hence, $lnB/\alpha < M_0 < \infty$.

Furthermore using Eq. 6.5

$$\frac{T_0 \cdot c_0}{V_0 \cdot I_0} = \frac{1 - P_0}{P_0} e^{\alpha M_0}$$

or

$$c_0 = \frac{V_0 \cdot I_0 \cdot (1 - P_0) e^{\alpha M_0}}{T_0 \cdot P_0}. \tag{6.6}$$

where $|(1 - P_0)/P_0| < 1$. No other requirements on V_0, T_0, I_0 or M_0 are present. Clearly, from our plots in section 6.4, curves with $A > 0.5$ possess the desired convexity changes which may for some values of the parameter produce curves which model a step function.

For a given region with known values of P_0, M_0 and α (for which $A/2 < P_0 < A$) one can find B_0 and subsequently, c_0. In addition, we can verify that,

$$\frac{1}{2} < P_0 < 1 \quad \text{for} \quad A = 1$$

$$\frac{1}{2} < \frac{1}{1 + B_0 e^{-\alpha M_0}} < 1$$

$$2 > 1 + B_0 e^{-\alpha M_0} > 1$$

or

$$1 > B_0 e^{-\alpha M_0} > 0$$

Hence, it is important to check that

$$e^{\alpha M_0} > B_0 > 0$$

for any estimation of P_0 using a numerical algorithm.

6.8 Solving for the lesion impulse density function

Recall that the density function for lesion impulses can be obtained by

$$\mathcal{L}^{-1}\left\{\frac{A}{1+Be^{-\alpha s}}\right\}$$

using our power series expansion.

Since

$$F(s_0) = \frac{A}{1+Be^{-\alpha s_0}}$$

we can write

$$F(s_0) = \lim_{N\to\infty}\sum_{k=0}^{N} A(-Be^{-\alpha s_0})^k \quad \text{for} \quad \frac{A}{2} \leq F(s_0) \leq A$$

where $\alpha > 0$. B is positive and real. We may write explicitly and in general.

$$F(s) = A\left[1 - Be^{-\alpha s} + B^2e^{-2\alpha s} - \cdots + B^n e^{-n\alpha s}\cdots\right] \text{ for } s = s_0$$

Recall that $F(s) \approx A$ (i.e., a step function) and applying the following Laplace Transform property,

$$F(s)e^{-\alpha s} \rightarrow f(t-\alpha)$$

one obtains

$$f(t) = A\left[\delta(t) - B\delta(t-\alpha) + B^2\delta(t-2\alpha) - \cdots + B^n\delta(t-n\alpha)\cdots\right]$$

or

$$f(t) = \sum_{n=0}^{\infty}(-1)^n AB^n\delta(t-n\alpha).$$

which represents a sequence of lesion impulses that reflect the sudden increase in lesion probability as depicted by our step function.

In the following figures. the function $F(s)$, and the Laplace Inverse of $F(s)$, that is $f(t)$, are shown respectively.

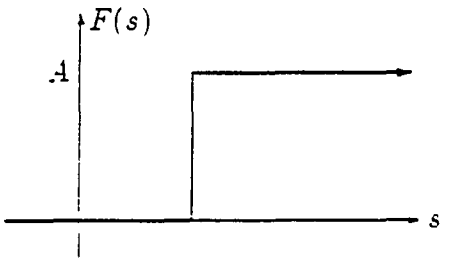


Figure 6.18: The Step Function $F(s)$

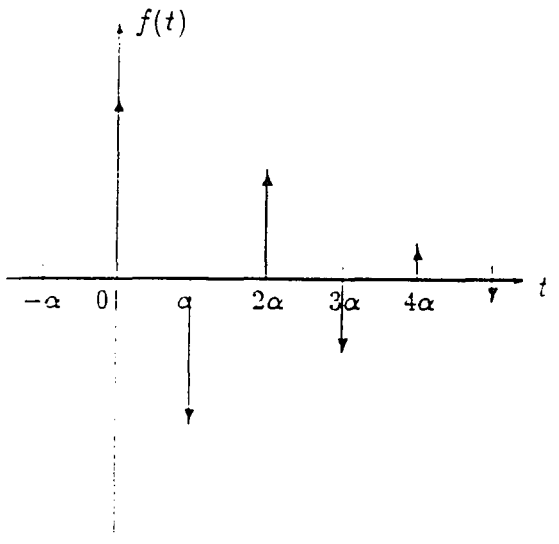


Figure 6.19: The Sequence of Lesion Impulses

Chapter 7

CONCLUSIONS

7.1 Summary

This thesis explored the localization of lesions in the human nervous system based on observed test outcomes. First, we addressed the issue of probability of lesion and turned to a Bayesian model as a technique for using á priori information to estimate posterior probabilities of lesion based on test outcomes. The Bayesian model was presented, with all its shortcomings, and hence we were led to explore certain methods and alternative models to deal with the sparsity of data and a limited neurological information base.

After a brief introduction, in the First Chapter, the literature survey was interposed focusing on papers and books dealing with relevant topics from 1959 to 1989. Chapter 2 was devoted to a Bayesian formulation of the CALOND. Eq. 2.8 of this chapter was the fundamental equation used for our preliminary studies. The purpose was to find the estimate of a lesion probability at a certain voxel for a given set of observed test outcomes. But the need to be able to estimate certain parameters of the fundamental equation was the main problem. For example, it was not feasible to obtain the á priori probability estimates for the presence of a lesion at a certain

voxel. Recall that, there are 1000 voxels, as defined in this model.

Observing the linear relation of the parameters of the fundamental equation, it was possible to apply a Regression Analysis, which in turn helped us to derive a functional relation between the unknown probability estimates. Initially, a linear first order model (with a physician's experience factor taken into account) was introduced. The lack of data was a major drawback which led us to simulation.

In Chapter 4, The Monte Carlo simulation was introduced: first the functional relation between the positive net-malfunction factor and negative net-malfunction factors were presented. Then, a new probability model was defined as:

$$P_s = c_s \frac{T_s \cdot M_s}{V_s \cdot I_s}$$

where, c_s was defined as the Calibration factor and I_s was defined as the Involvement factor. On the basis of this model, for a specific case–Friedreich's Ataxia–the simulation results showed that the calibration factors hardly differ if the given brainstem section contains a rather dense lesion distribution, and if the distribution contains lack of geometrical connectivity, then there was a notable difference between the real and simulated calibration factors.

The descriptive statistics of the simulation were studied in Chapter 5. The mode, mean and the median were discussed in relation to the probability model posed for CALOND, since they served as the basis for subregion determination in our simulation.

The Logistic Sigmoid Model was introduced in Chapter 6. The model was defined as:

$$P_s = \frac{A}{1 + B_s e^{-\alpha M_s}}$$

and the relation between this model with the linear one was depicted in the value of B as:

$$B_s = c_s \frac{T_s}{V_s \cdot I_s}$$

Under these assumptions, the plots of the probability of lesion versus the Positive Net-Malfunction Factor, for different values of α and P_s, were compared and it was shown that $0.1 < \alpha < 0.5$ and $A > 0.5$ provide the expected convexity of the curves in our model. In the last portion of this Chapter, a severely convex logistic sigmoid model, i.e., a step function–like approach, was analyzed. It was shown that, the distribution for lesion probability, $P(M) = A/(1 + Be^{-\alpha M})$, may be viewed as the moment generating function of a lesion impulse density, which for large α approaches a step function. Using transform methods a sequence of lesion impulses, which generated this large jump in lesion probability, was derived.

7.2 Conclusions

From the preceding summary, we focus upon some salient conclusions:

1. Modelling a predictor probability distribution to localize neurological lesions may take a variety of approaches and formulations–from strictly empirical estimation–to randomized simulation of patient and disease attributes–to more deterministic perspectives which identify a classical case (such as logistic sigmoid nonlinearity) and then parameterize it to appropriately reflect true characteristics of the initial problem under study.

2. Computer aided medical diagnosis, which makes use of the CALOND database, must find a medium through which the physician may associate an underlying probability of lesion distribution with the patient currently undergoing a neurological examination. This connective medium must consolidate patient attributes and symptomology (in the form of neurological test outcomes) and apply this á priori information to a synthesizing process which produces a posterior probability of lesion distribution. The groundwork for this synthesis has been laid in the following discussions.

a) The initial step in the generation of a posterior lesion probability distribution involves the application of Bayes' Theorem. An appropriate statistical process has been identified to estimate posterior probabilities: regression analysis. This process is inherently compatible with our goals because of the existing linearity mandated by Bayes' Theorem.

b) A primary contribution made by this thesis is the subsequent identification of appropriate linear and non–linear models which incorporate personal and symptomatic attributes of a patient undergoing a neurological examination into the Bayesian (i.e., conditional: based upon all available historic information) framework.

c) Randomized simulation is applied to extract an estimate of the true underlying probability of lesion in cases where empirical data cannot be feasibly obtained. This simulation associates underlying patient attributes with neurological test outcomes to produce posterior lesion probabilities.

d) Techniques for converging upon anatomical subregions which contain the centroid of a lesion (or set of lesions) are presented and compared by way of a simulation for a specific example. In addition, overall descriptive statistics of net malfunction factors are used to set guidelines for identifying anatomical regions of function, malfunction and ambiguity.

e) The thrust of this work has been to consolidate and interrelate geometrical, algebraic, statistical, and probabilistic elements of CALOND–which exists as a database in the form of pathways of function or malfunction–and to use this interrelatedness to best localize neurological lesions.

f) The concept and potential existence of net–malfunction factor threshold values were explored. These threshold values indicate levels of criticality present in the neurological examination and are understood to reflect a certain "risk" value which exists for patients with the specific set of per-

sonal and symptomatic attributes which produce those values of certain parameters in the "at–risk" intervals

3. This thesis lays the groundwork for some useful and much needed future research as presented in the following section.

7.3 Suggestions for Future Research

The author believes that, the work done in this thesis was indeed a very small unit of possible research topics which would be appropriate follow up studies. The following are areas for potential future research:

- Use the results of the CAT Scanning or NMR in combination with CALOND, to find the probable lesion locations.
- Expand CALOND to layout the pictures of the section under study–even though the colored sections are prepared, they have yet to be used.
- Develop a user friendly, marketable product.
- Explore neurological diseases which possess a probability distribution which approaches a step function and identify the threshold value for M (i.e., the inflection point).
- Associate specific values of α, B, A to certain diseases using nonlinear regression analysis.
- Explore additional techniques for the determination of subsequent subregions in the Monte Carlo Simulation, other than the measures of centrality used in this thesis. This would aid in the analysis of lesions that are not spatially connected.

- Explore differing underlying patient population distributions. such as Bernoulli, Gamma, Poisson, Cauchy, etc., to characterize the role of individual human attributes and how they relate to neurological test outcomes.
- Explore neurological examination procedures and isolate an appropriate order for test application based on some initial selective criteria.

Chapter 8

Glossary

Symbol	Range and Definitions	Page
b_i;	Estimated values of the parameters β_i	113
c_s;	Calibration Factor	40
$f(M)$;	Discrete probability of malfunction for any subset of a fundamental set of voxels	51
F;	Function Factor	36
$F(s)$;	Double sided Laplace Transform Function	92
F_s;	Sum of the negative NMF in section s	36
F_v;	$(v = 1, \ldots, V)$ Event that voxel v is functioning	21
i;	$(i = 1, \ldots, I)$ Identification number of test (listed) in the TESTBASE.	21
$i(n)$;	$(n = 1, \ldots, N)$ ID number of nth test performed	21
I_s;	Involvement Factor in section s	39
j;	$(j = 1, \ldots, J_I)$ Identification number of outcome (listed) in the TESTBASE for test i	21
$j(n)$;	$(n = 1, \ldots, N)$ ID number of outcome of nth test performed	21
k_s;	Inverse Malfunction Constant	51
M;	Malfunction Factor	36
$M_X(\Theta)$;	Moment Generating Function	92
M_s;	Sum of the positive NMF in section s	36
M_v;	$(v = 1, \ldots, V)$ Event that voxel v is malfunctioning	21
n;	$(n = 1, \ldots, N)$ (Consecutive) test sequence number	21
N;	Total number of tests performed	21
NMF;	Net Malfunction Factor	36

Symbol	Range and Definitions	Page
$O_{i,j}$;	$(i = 1, \ldots, I;\ j = 1, \ldots, J)$	
	Verbal description of jth outcome of test i	21
P_s;	The probability of malfunction in the	
	section s for a given set of tests outcomes	51
$\mathcal{O}$;	$\mathcal{O} = (O_{i,j};\ i = 1, \ldots, I;\ j = 1, \ldots, J_I)$	
	Set of all test outcomes	21
$\boldsymbol{q}(n)$;	$\boldsymbol{q}(n) = \begin{bmatrix} i(n) \\ j(n) \end{bmatrix} \quad (n = 1, \ldots, N)$	
	Test outcome (vector) for nth test performed	21
R_s;	Randomly chosen region s under study	51
$\mathcal{S}$;	Residual sum of Squares	28
$\mathcal{T}$;	$\mathcal{T} = (T_1, \ldots, T_I)$	
	Set of all tests in CALOND	21
T_i;	$(i = 1, \ldots, I)$	
	Verbal description for test i	21
T_s;	Total number of tests applied in	
	section s	40
v;	$(v = 1, \ldots, V)$	
	ID number of voxel v	19
V_s;	Total number of voxels involved in	
	section s	40
$\boldsymbol{V}(x(y))$;	Variance of the distance $x(y)$	57
$\boldsymbol{V}(\boldsymbol{b})$;	Variance, covariance matrix for	
	Regression coefficient matrix b	29
$\bar{X}(Y)$;	distance of voxel with respect to origin	
	$(0, 0)$ with respect to X(Y)	57
β_i;	$(i = 0, 1)$	
	Unknown parameters of the linear regression	
	model	27
$\mu_{x(or\ y)}$;	Mean of the distance in the $x(y)$	
	direction	57
θ_s;	Malfunction angle	36

Appendix A

Brainstem

The following material is excerpted from " Human and Physiology-Structure and Function" by D. S. Luciano, et. al., McGraw Hill, Second Ed., 1983.

Brainstem is literally the stalk of the brain, through which passes all the nerve fibers that relay signals of afferent input and efferent output between the spinal cord and higher brain centers. In addition, the brainstem contains the cell bodies of neurons whose axons go out to the periphery to innervate the muscles and glands of the head, the heart, and the smooth muscles and glands of most thoracic and abdominal viscera. The brainstem also receives many afferent fibers from the head and visceral cavities via the cranial nerves. In contrast to the distinct white and gray areas of the spinal cord, the tracts and nuclei of the brains are intermingled.

The **medulla oblongata** is the section of the brainstem continuous with the spinal cord below and the pons above. Its junction with the cord reflects a gradual change from the external tracts and internal columns of nuclei that exits at the upper levels of the cord. Efferent axons emerging from the medulla via cranial nerves VIII, IX, X, XI, and XII control areas of mouth, throat, neck, throax, and abdomen.

The **pons** is both wider and thicker then the medulla and is easily distinguished by a band of fibers running across its ventral surface. These fibers converge at each side of the pons into bundles called the **middle cerebellar peduncles**, one of

the three pairs of fiber bundles that carry information between the brainstem and cerebellum. The afferent and efferent components of cranial nerves V, VI, and VII connecting with the pons are from the head.

The **midbrain** is a relatively short part of the brainstem and is somewhat constricted in comparison with the pons. It is traversed by a huge number of axons that contribute to the corticospinal and spinocortical pathways. It contains major nuclei associated with eye movements and hearing.

Running through the entire brainstem is a core of tissue called the **reticular formation**, which is composed of a diffuse collection of small, many branched neurons. The neurons of the reticular formation receive and integrate information from many afferent pathways as well as from many other regions of the brain. Some reticular formation neurons are assembled together, forming certain of the brainstem nuclei and "centers", such as the cardiovascular, respiratory, swallowing, and vomiting centers. The output of the reticular formation can be divided functionally into descending and ascending systems. The descending components influence efferent neurons in the cranial and spinal nerves and frequently afferent neurons as well; the ascending components affect such things as wakefulness and the direction of attention to specific events.[10]

A.1 Definition of selected neurological terms

afferent pathway: The components of a reflex arc that transmits information from a receptor to an integrating center: any pathway that conveys information toward the central nervous system (or toward the brain).

corticospinal pathway: A descending motor pathway that has its nerve cell bodies of origin in the cerebral cortex; the axons pass without synapsing to the region of the motor neurons; also called *pyramidal tract.*

efferent pathway: That component of a reflex arc that transmits information

from the integrating center to the effector; any pathway that conveys information out of the central nervous system (or away from the brain within the central nervous system).

ventral root: A group of efferent fibers that leaves the left and right side of the region of the spinal cord that faces the front of the body.

viscera: The organs in the thoracic and abdominal cavities

Appendix B

A Regression Analysis Model

The following material is excerpted from " Applied Regression Analysis", by N. R. Draper and H. Smith, Wiley and Sons, 1988.

For any system of which variable quantities change, it is possible to examine the functional relations between variables. Often these functional relations might be too complicated to handle or describe in simple way. But it is possible to approximate these relationships by some simple mathematical function. Then, it will be possible to learn more about these variables. In this research, a linear relation in unknown parameters is assumed. These unknown parameters are estimated under certain other assumptions with the aid of the available data. In this case it is the construction of a fitted straight line with the pairs of observations $(X_1, Y_1), \cdots, (X_n, Y_n)$ and the model is linear, first-order, given as;

$$Y = \beta_0 + \beta_1 X + \epsilon$$

where X is the given data, Y is the corresponding observation, ϵ is the increment by which any individual Y may fall off the regression line, or simply the error. β_0 and β_1 are called the *parameters* of the model. The estimates of these parameters, b_0 and b_1 can be find as follows:

$$\hat{Y} = b_0 + b_1 X$$

where $\hat{Y}$ is the predicted value of Y for a given X. For an estimation *Least Squares* procedure will be applied that is:

$$S = \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_i)^2$$

which shows the sum of the squares of deviations from the true line. After certain steps of mathematical calculations the following will be derived for b_1 and b_0 respectively;

$$b_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{X})}{\sum(X_i - \bar{X})^2}$$

$$b_0 = \bar{Y} - b_1 \bar{X}$$

Knowing b_1 and b_0, it will be possible to find the confidence interval for β_1 and β_0.

Let $\sum(X_i - \bar{X})Y = \bar{Y}\sum(X_i - \bar{X}) = 0$, than the equation for b_1 will be as follows:

$$b_1 = \frac{\sum(X_i - \bar{X})Y_i}{\sum(X_i - \bar{X})^2}$$

the variance for b_1 will be:

$$Var(b_1) = \frac{\sigma^2}{\sum(X_i - \bar{X})^2}$$

and the standard deviation will be:

$$s.e(b_1) = \frac{\sigma}{\{\sum(X_i - \bar{X})^2\}^{1/2}}$$

for σ known.

Under the assumption that variation of the observations about the line are normal, that is, the errors ϵ_i are all from the same normal distribution, $N(0, \sigma^2)$, then $100(1 - \sigma)\%$ confidence limits can be assigned for b_1 by calculating,

$$b_1 \pm \frac{t(n-2, 1 - \frac{1}{2}\alpha)\sigma}{\{\sum(X_i - \bar{X})^2\}^{1/2}}$$

where $t(n-2, 1 - \frac{1}{2}\alpha)$ is the $(1 - \frac{1}{2}\alpha)$ percentage point of a t-distribution, with $(n-2)$ degrees of freedom.

In a similar way, a confidence interval for β_0 and the test of whether or not β_0 is equal to some specified value, can also be constructed. In this case.

$$s.e(b_0) = \left\{\frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}\right\}^{1/2}\sigma$$

thus $100(1-\alpha)\%$ confidence limits for β_0 are given by,

$$b_0 \pm t(n-2, 1-\frac{1}{2}\alpha)\left\{\frac{\sum X_i^2}{n\sum(X_i - \bar{X})^2}\right\}^{1/2} s$$

Finally the standard error of $\hat{Y}$ can be calculated as follows, first the variance of the predicted mean value of Y, $\hat{Y}_k$ at a specified X_k, of X is,

$$V(\hat{Y}_k) = V(\hat{Y}) + (X_k - \bar{X})^2 V(b_1) = \frac{\sigma^2}{n} + \frac{(X_k - \bar{X})^2\sigma^2}{\sum(X_i - \bar{X})^2}$$

which gives

$$est.s.e(\hat{Y}_k) = s\left\{\frac{1}{n} - \frac{(X_k - \bar{X})^2}{\sum(X_i - \bar{X})^2}\right\}^{1/2}$$

which is minimum when $X_k = \bar{X}$ and decreases as X_k moves away from $\bar{X}$ in either direction. In other words, the greater distance an X_k is from $\bar{X}$, the larger the error expected to get. [9]

Appendix C

Monte Carlo Simulation

The following material is excerpted from " Simulation and the Monte Carlo Method" by R. Y. Rubinstein. Wiley and Sons, 1981

The Monte Carlo method-or method of statistical trials- consists of solving various problems of computational mathematics by means of the construction of some random process for each problem, with the parameters of the process equal to the required quantities of the problem. These quantities are then determined approximately by means of observations of the random process and the computation of its statistical characteristics, which are approximately equal to the required parameters.

In the more strict sense of the term, the Monte Carlo method is defined as the construction of an artificial random process possessing all the necessary properties, but which is in principle realizable by means of ordinary computational tools.

The following situations will show where the simulation can be used successfully.

(1) It may be impossible or extremely expensive to obtain data from certain processes in the real world. Thus simulated data are necessary to *formulate hypothesis* about the system.

(2) The observed system may be so complex that it can not be described in terms of a set of mathematical equations for which analytic solutions are available.

(3) Even though a mathematical model can be formulated to describe some

system of interest, it may not be possible to obtain a solution to the model by straight forward analytic technique.

Although it may be conceptually possible to use set of mathematical equations to describe the behavior of a dynamic system operating under conditions of uncertainty present-day mathematics and computer technique are simply incapable of handling a problem of this magnitude.

(4) It may be either impossible or very costly to perform validating experiments on the mathematical models describing the system, thus the simulated data can be used to test alternative hypotheses.

Simulation analysis might be appropriate for the following reasons.

(1) Simulation makes it possible to study and experiment with the complex internal interactions of a given system.

(2) One can study the effects of certain informational, organizational, and environmental changes on the operation of a system by making alterations in the model of the system and observing the effects of these alterations on the system's behavior.

(3) Detailed observation of the system being simulated may lead to a better understanding of the system and to suggestions for improving it, suggestions that otherwise would not be apparent.

(4) Simulation can be used as a pedagogical device for teaching both students and practitioners basic skills in theoretical analysis, statistical analysis, and decision making. Among the disciplines in which simulation has been used successfully for this purpose are business administration, economics, medicine, and law.

(5) The knowledge obtained in designing a simulation study frequently suggests changes in the system being studied. The effects of these changes can then be tested via simulation before implementing them on the actual system.

(6) Simulation of complex systems can yield valuable insight into which variables are more important than others in the system and how these variables interact.

(7) It can be used to experiment with new situations about which we have little or no information so as to prepare for what may happen.

(8) It can serve as a *preservice test* to try out new policies and decision rules for operating a system, before running the risk of experimenting on the real system.

(9) They are sometimes valuable in that they afford a convenient way of breaking down a complicated system into subsystems, each of which may then be modeled by an analyst or team that is expert in that area.

(10) It makes it possible to study dynamic systems in either real time, compressed time, or expanded time.

(11) When new components are introduced into a system, simulation can be used to help forsee bottlenecks and other problems that may arise in the operation of the system.

Computer simulation allows us to induce correlation between the random number sequences to improve the statistical analysis of the output of a simulation. In particular a negative correlation is desirable when the results of two replications are to be summed, whereas a positive correlation is preferred when the results are to be differenced.

Simulation dose not require that a model be presented in a particular format. It permits a considerable degree of freedom so that a model can bear a close correspondence to the system being studied.

Simulation is by no means ideal. It is an imprecise technique. It provides only a *statistical estimates* rather then exact results, and it only compares alternatives rather then generating the optimal one. It is also a "slow" and "costly" way to study a problem. It yields only *numerical data* about the performance of the system, and sensitivity analysis of the model parameters is very expensive. The only possibility is that to conduct series of simulation runs with different parameter values

Simulation is defined as a technique of performing sampling experiments on

the model of the system. Since sampling for a particular distribution involves the use of random numbers, stochastic simulation is sometimes called *Monte Carlo Simulation*. The Monte Carlo Method was considered to be a technique, using random or pseudorandom numbers, for solution of a model. Random numbers are essentially independent random variables uniformly distributed over the unit interval [0,1]. [28].

Appendix D

List of tests/outcomes used in CALOND

Test 1: Observe vocal cords during phonation

Left cord weak or paralyzed, right motion normal

Right cord weak or paralyzed, left motion normal

Test 2: Ask patient to say "ah", observe oropharynx

Weak or paralyzed left palate

Normal Palatal action

Weak or paralyzed right palate

Bilateral Palatal Weakness or Paralysis

Test 3: Observe pupillary size

Left constricted, right larger

Right constricted, left larger

Test 4: Pin-prick applied to right limbs and right torso

No pain experienced

Pain acutely experienced

Test 5: Observe pupillary size

Right redilates

Left redilates

Test 6: Shine light in right pupil

Right contracts markedly, Left contracts minimally

Test 7: Observe pupils

Right redilates

Left remains miotic

Test 8: Shine light in left pupil

Right contracts markedly, Left contracts minimally

Test 9: Ask patient to look up to right, then up to left

Movements normal

Test 10: Ask patient to look down to right, then down to left

Movements normal

Test 11: Ask patient to follow the moving finger horizontally

Nystagmus present

Eyes track normally

Test 12: Touch patient's forehead to estimate moisture

Right moist, left dry

Test 13: Pin-prick applied to left limbs and left torso

Pain acutely experienced

Test 14: Ask patient to follow rising finger

Right lid elevates, left drops

Test 15: Have patient alternately touch each index finger to nose

Right motion normal, Left motion awkward

Test 16: Have patient slide each heel along opposite shin

Right motion normal, Left awkward

Test 17: Touch right corneal edge with cotton whisp

Eyelids blink

Right lid blinks only

Test 18: Touch left corneal edge with cotton whisp

Neither eyelid contracts

Test 19: Place sweet, sour or salt solution on each anterior half tongue

Right savors, Left impaired

Test 20: Place sweet, sour or salt solution on each posterior third tongue

Right savors, Left does not

Test 21: Move right thumb, distal phalanx to test motion sense

Proprioception normal

Test 22: Move left thumb, distal phalanx to test motion sense

Proprioception normal

Test 23: Move right great toe, distal phalanx, to test motion sense

Proprioception normal

Test 24: Move left great toe, distal phalanx, to test motion sense

Proprioception normal

Test 25: Observe vocal cords while patient breathes quietly

Normal bilateral cord activity

Test 26: Whisper words in patient's right ear

Patient repeats word properly

Test 27: Whisper words in patient's left ear

Patient repeats word properly

Patient denies hearing

Test 28: Touch to irritate right cornea

Tearing induced on right

Test 29: Touch to irritate left cornea

No tearing on left

Tearing induced on left

Test 30: Whisper words in patient's left ear

Patient denies hearing

Patient hears properly

Test 31: Perform right caloric test (cold water)

Horizontal nystagmus, slow phase to right

Test 32: Perform left caloric test (cold water)

Horizontal nystagmus, slow phase to left

No nystagmus

Test 33: Observe vocal cords attempting phonation

Cords remain adducted

Normal bilateral cord excursion

Test 34: Touch right pharyngeal wall

Patinet feels touch

Patient can not feel touch

Test 35: Touch left pharyngeal wall

Patient feels touch

Patient can not feel touch

Test 36: Stimulate right posterior pharyngeal wall

Normal gag reflex induced

Gag reflex not elicited

Test 37: Stimulate left posterior pharyngeal wall

Gag reflex not elicited

Normal gag reflex induced

Test 38: Observe tongue in patient's mouth

Fasciculations/atrophy on right only

Test 39: Hot/cold application to left chin

Thermal sense perceived

Test 40: Ask patient to demonstrate chewing ability

Jaw motion normal

Test 41: Tap chin to elicit jaw-jerk

Normal reflex

Test 42: Pin-prick applied to right forehead

Patient does not feel pain

Test 43: Pin-prick applied to left forehead

Patient experiences pain

Test 44: Pin-prick applied to right cheek

Patient experiences pain

Test 45: Pin-prick applied to left cheek

Patient experiences pain

Test 46: Pin-prick applied to right chin

Patient does not feel pain

Test 47: Pin-prick applied to left chin

Patient experiences pain

Test 48: Light touch applied to right forehead

Touch perceived

Test 49: Light touch applied to left forehead

Touch perceived

Test 50: Observe protrusion of patient's atrophic tongue

Tongue deviates to right

Test 51: Ask patien to wrinkle brow

Left forehead furrows. right flat

Test 52: Ask patient to squeeze eyelids shut

Left normal, right closure incomplete

Test 53: Ask patient to exaggerate smile

Left lips active. right weak with flat or sagged aspect

Appendix E

Voxel numbering and location identification

In the example shown on the next page, 3 of the 10 sections are displayed with corresponding x and y coordinates. On the vertical coordinate the numbers are shown as $i0, i1, \ldots, i9$, where i refers to section number. Also shown are some numbers to indicate the malfunction factors. For example, $v = 1022$, is the voxel in section 10 with $x = 2$ and $y = 2$, then $M(v) = M(1022) = 4$, that is, voxel v has a malfunction factor of 4, for $v = 272$, the voxel is in section 2 with $x = 2$ and $y = 7$ and $M(v) = M(272) = 3$. In CALOND there are $v = 1000$ voxels, $v = 100, 101, \ldots, 1099$.

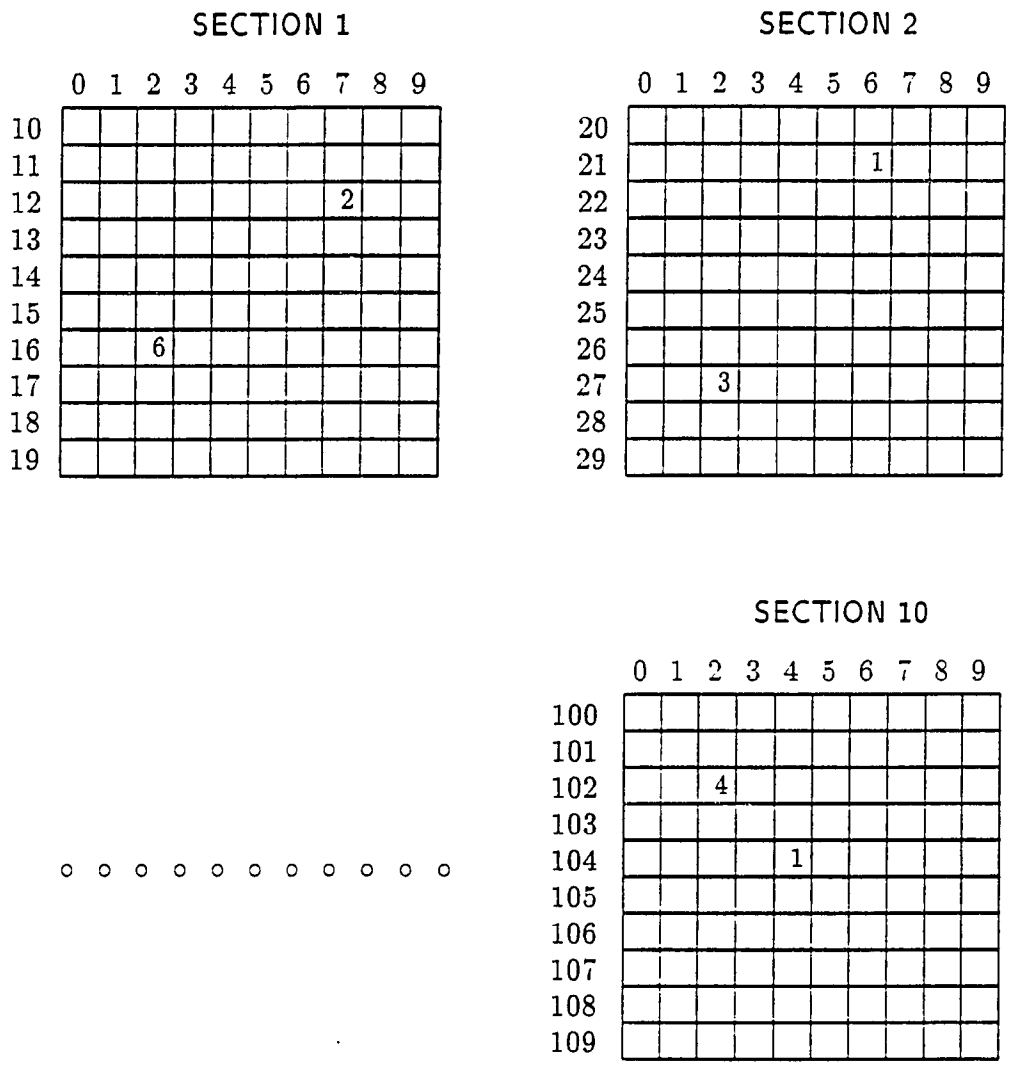


Figure E.1: Display of sections and voxel, v representation.

Bibliography

[1] Adlassnig, K. P. , et. al.,"CADIAG:Approaches to Computer–Assisted Medical Diagnosis,"*Comput. Biol. Med.* , Vol. 15 No. 5, 1985, pp. 315–335.

[2] _____ ," Representation and Semiautomatic Acquisition of Medical Knowledge in CADIAG–1 and CADIAG–2," *Comp. Biomed. Res.* ,Vol. 19, 1986, pp. 63–79.

[3] _____." Performance evaluation of medical expert systems using ROC curves".*Comp. Biomed. Res.*, Vol. 22, 1989, pp. 297–313.

[4] Ben–Bassat, Moshe, et. al.," Pattern–Based Interactive Diagnosis of Multiple Disorders: The MEDAS system. "*IEEE Trans. Pattern Analysis Mach. Intell.* , Vol. PAMI–2, No. 3, 1980, pp. 148–160.

[5] *Blakiston's Illustrated Pocket Medical Dictionary*, The Blakiston Division, McGraw Hill Book Com. , 1960, Second Edition.

[6] Bleich, H. L. ,"Computer Evaluation of Acid-Base Disorders," *Journal of Clinical Investigation*, Vol. 48, 1969, pp. 1689–1696.

[7] Catanzarite, V. A. "NEUROLOGIST : A Computer Program for Localization and Diagnosis in Clinical Neurology," Ph. D. Thesis, University of California, Berkeley, 1980.

[8] "*Clinical Neurology–The Epidemiology of Neurologic Disease*", Vol. 4, Chapter 66, J. B. Lippincott Com. , 1989.

[9] Draper, N. R. and Smith, H. ,"*Applied Regression Analysis*," Wiley and Sons, 1988.

[10] Luciano, D. S. , Vander A. J. , Shermen, J. H. , "*Human and Physiology–Structure and Function*", McGraw Hill, Second Ed., 1983.

[11] Du Boulay, G. H. et. al. ,"The Diagnosis of Intracranial Tumours Assisted by Computers, "*British Journal of Radiology*, Vol. 41, No. 490, 1968, pp. 762–781.

[12] _____,"Improvement in the Computer–Assisted Diagnosis of Cerebral Tumours,"*British Journal of Radiology*, Vol. 50, No. 600, 1977, pp. 849–854.

[13] First, M. B. et. al. ,"LOCALIZE: Computer–Assisted Localization of Peripheral Nervous System Lesions, "*Comput. Biomed. Res.* ,Vol. 15, 1982, pp. 525–543.

[14] Gorry, G. A. and Barnett, G. O. ,"Experience with a Model of Sequential Diagnosis,"*Comput. Biomed. Res.* , Vol. 1, 1968, pp.490–507.

[15] _____ ,"Decision Analysis as the Basis for Computer-Aided Management of Acute Renal Failure,"*Amer. J. Med.* , Vol. 55, 1973, pp. 473–484.

[16] "*Greenfield's Neuropathology*", Edited by J. H. Adams, J. A. N. Corsellis and L. W. Duchen, 4th Edition, John Wiley & Sons, Inc., N.Y, 1984

[17] Knapp, R. G. et. al. ,"A computer–generated diagnostic decision guide: A comparison of statistical diagnosis and clinical diagnosis," *Comput. Biol. Med.* , Vol. 7, 1977, pp.223–230.

[18] Ledley, R. S. and Lusted, L. B. ,"Reasoning foundations of medical diagnosis," *Science*, Vol. 130, 1959, pp. 9–21.

[19] Meyer, A. U. and Weissman, W. K. ,"Localization of lesions in the human nervous system by computer analysis,"*IEEE Trans. Biomed. Eng.* , Vol. BME–20, 1973, pp. 194-200.

[20] _____,"Computer Analysis of the clinical neurological examination" *Comput. Biol. Med.*, 1973.

[21] Mori, H. , et. al. ,"Differential diagnosis of brain lesions with a computed brain scan diagnosis by the likelihood method," *Investigative Radiology*, Vol. 10, 1975, pp. 251–257.

[22] Ohmann, C. , et. al. ,"Bayes theorem and conditional dependence of symptoms: Different models applied to data of Upper Gastrointestinal bleeding",*Methods of Inf. in Med.*, Vol. 27, 1988, pp. 73–83.

[23] Okada, M. , et. al. ,"Medical data base system with an ability of automated diagnosis," *Compt. Prog. Biomed.* , Vol. 7, 1977, pp. 163–170.

[24] Pople, H. ,"INTERNIST–1:An experimental computer–based diagnostic consultant for general internal medicine."*New England J. Med.* ,Vol. 307, No. 8, 1982, pp.468–476.

[25] Reggia, J. A. ," A production rule system for neurological localization,"In 'Proceedings, second annual symposium on computer applications in medical care'*IEEE Compt. Soc.* , F. Orthner, (ed.), Washington D.C, 1978, pp. 254–260.

[26] Rubinstein, R. Y. ,"*Simulation and the Monte Carlo Method*," Wiley and Sons, 1981.

[27] Shortliffe, E. H. ,"*Computer–Based Medical Consultations: MYCIN*,"Elsevier, New York, Oxford, 1976.

[28] "*The Monte Carlo Method*," Edited by Yv. A. Shreider, Pergamon Press, 1966.

[29] Sobol, I. M. "*The Monte Carlo Method*", The University of Chicago Press, 1974.

[30] Stewart, A. J. and Cala, L. A. ,"Mathematical method to utilize a computer diagnosis of site and type of intracerebral mass lesions,"*British Journal of Radiology*, Vol. 48, 1975, pp. 97-100.

[31] Walpole, R. E. and Myers, R. H.,"*Probability and Statistics for Engineers and Scientists*", 4th Edition, MacMillan, 1989.

[32] Warner, H. R. et. al. ,"Experience with Bayes' Theorem for computer diagnosis of congenital heart disease."*Ann. N.Y. Acad. Sci.*, Vol. 115, 1964, pp. 558-567.

[33] Weiss, N. A. and Hasset, M. J. ,"*Introductory Statistics*", 2nd. Ed. Addison Wesley, 1987.

[34] Wiener, Fred,"Computer simulation of diagnostic process in medicine," *Comput. Biomed. Res.*, Vol. 8, 1975, pp. 129–142.

[35] Wortman, P. M. ,"Medical Diagnosis: An information–Processing Approach," *Comput. Biomed. Res.* , Vol. 5, 1972, pp. 315–328.

[36] Xiang, Z, et. al. ,"Computerized Neurological Diagnosis : A paradigm of modeling and reasoning," *Health Care Instrumentation*, Vol. 1, 1986, pp. 90–105.