MIGUEL ANGEL GUZMAN

STUDY AND APPLICATION OF A NON-LINEAR MODEL FOR THE NUTRITIONAL EVALUATION OF PROTEINS

STUDY AND APPLICATION OF A NON-LINEAR MODEL FOR THE NUTRITIONAL EVALUATION OF PROTEINS

by

MIGUEL ANGEL GUZMAN

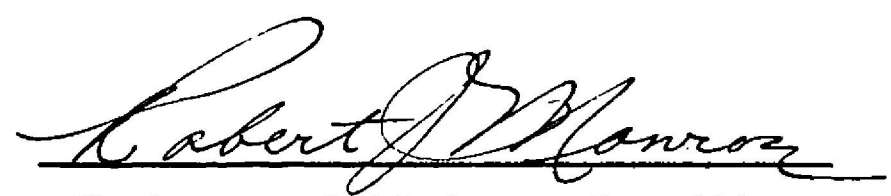
A thesis submitted to the Graduate Faculty
of North Carolina State College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

DEPARTMENT OF EXPERIMENTAL STATISTICS

RALEIGH

1961

APPROVED BY:



Chairman of Advisory Committee

T O

TOYA, MIKE, DICK AND TANI

BIOGRAPHY

The author was born to Patrocinio Guzmán Trigueros and Elsa Foresti de Guzmán Trigueros on December 20, 1926, in San Vicente, El Salvador, Central America. He initiated his high school education at the "Liceo Salvadoreño" in San Salvador, completing it in Memphis, Tennessee, at "Christian Brothers College". In 1949 the author received the degree of Bachelor of Arts in Zoology and Chemistry from the University of Tennessee; from 1949 to 1954 he was employed as a biochemist by the "Instituto de Nutrición de Centro América y Panamá" (INCAP) serving this institution successively as Biochemist, chief of the Section of Microanalysis and chief of the Laboratories. In 1954 the author was granted study leave by INCAP for graduate work in Experimental Statistics at North Carolina State College; during this period he was an International Fellow of the W. K. Kellogg Foundation from September 1954 to August 1956, and a Rockefeller Foundation Fellow from September 1956 to June 1957. He received the degree of Master of Science in Experimental Statistics from North Carolina State College in January 1957. Upon return to INCAP in June 1957, the author was appointed chief of the Division of Statistics, and has served INCAP to date in this capacity.

On different occasions, the author has served as a short term statistical consultant for the World Health Organization of the United Nations (WHO), duPont de Nemours and the Department of Agriculture of the United States. The author is a member of the American Statistical Association, the Biometric Society and the American Institute of

Nutrition; he is also a member of the Society of Sigma Xi and has been elected to membership in the Honor Society of Phi Kappa Phi.

In December 1951, the author was married to María Victoria Sobalvarro and from this marriage three children have been born: Miguel Antonio on February 9, 1953; Ricardo Enrique on June 12, 1955, and Ana Cecilia on June 26, 1957.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude to Dr. Robert J. Monroe and the other members of my graduate committee: Drs. Gennard Matrone, Henry L. Lucas, Richard L. Anderson and Walter J. Peterson. Their continued encouragement, suggestions and assistance were always timely and of great value throughout the period of my studies and especially in the preparation of this dissertation.

I am greatly indebted to Dr. Nevin S. Scrimshaw for his continued interest and assistance which have made it possible that I reach this important stage in my career. To Professor Gertrude M. Cox I am grateful for directing my interest to the field of Applied Statistics. I am grateful also to my colleagues at INCAP for the confidence they have always deposited in me.

An expression of appreciation is also due to my classmates and fellow students, especially to Mr. G. T. Foradori, for stimulating discussions and pertinent suggestions through the period of these studies. The kindness of Mr. Robert Teichman in making available to me material from his reprint files is greatly appreciated. I am also indebted to Mrs. Edith King for technical assistance in the IBM laboratory, to Mrs. K. M. McEntire for typing the final copy of the manuscript, and to Miss Mary Ann Cipolloni for assistance in the verification of references and the proof-reading of the final copy.

Lastly, I am deeply grateful to my wife and children for their understanding and continued sacrifices throughout the years of study which represent an invaluable contribution to the realization of my educational and professional achievements.

This research was supported jointly by the Institute of Nutrition of Central America and Panama (INCAP), Guatemala, C. A., through a special grant from the Williams Waterman Research Corporation, and by the Institute of Statistics of the Consolidated University of North Carolina through grant RG-7143 from the National Institutes of Health (NIH).

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

Page

CHAPTER I

PROTEINS IN NUTRITION

1.1 Introduction

Poor nutrition is responsible for a large share of the total mortality in the world, choosing most of its victims among children under five years of age. Indeed, for some of the technically underdeveloped areas of the world, it has been estimated that 25-30% of the children born never reach the age of five years (Scrimshaw and Béhar, 1959, 1961). It has been suggested that most of these deaths, especially in the 1-4 year age group, are due principally to poor nutrition (Gómez, et al. 1958; Hundley, 1959, Guzmán, et al., 1961).

Fortunately, the problem of poor human nutrition has been recognized as a public health hazard and attention is being given to its solution (Sebrell and Hand, 1957; Perez, 1959). Progress has been made in demonstrating that certain diseases can be cured through the administration of proper diets. For example, "kwashiorkor", now recognized as a severe type of protein deficiency which occurs characteristically in underprivileged population groups amongst children aged 1-4 years, responds readily to treatment with milk or other suitable proteins (Brock and Autret, 1952; Béhar, et al., 1958). To emphasize the practical importance of protein in the general diets consumed in most technically underdeveloped areas, it should be mentioned that "kwashiorkor" and other forms of protein deficiency are primarily responsible for the high child mortality mentioned earlier (Scrimshaw and Béhar, 1959; 1961).

Although protein can be a limiting factor in a diet in terms of either quantity or quality, generally it is found limiting in both respects, and an actual deficiency is often conditioned by the relative amounts of other nutrients present in the diet. Thus, the quantitative and qualitative aspects of a diet interact in such a manner that it is not always possible to correct a conditional qualitative deficiency by increasing the quantity of protein in the diet.

The investigation of the fundamental parameters of protein utilization should be useful in devising better ways for deriving in practice the maximum benefits possible from the available protein supplies. Additionally, this information should facilitate the search for cheaper sources of protein and/or the formulation of combinations of different foods which can have a greater food value than any of its components through proper complementation of their nutritive elements.

1.2 Chemistry and General Metabolism of Proteins

Proteins occupy a very important place in both the structural and the dynamic aspects of all living matter. Nearly half of the dry matter of adult man is protein, and next to water this is the most abundant material in the animal body. About one third of the protein is in muscle, one fifth in bone and cartilage, one tenth in the skin; the rest is found in other tissues and body fluids, with the exception of bile and urine which do not normally contain protein. Proteins are also very specific in many of their functions in the body and cannot be replaced by other substances in their basic role in metabolism.

1.2.1 Historical. Recognition of the importance of the proteins in the biological functions occurred during the first half of the nineteenth century. Many investigators had noted that substances like egg white, blood serum and milk curd had many characteristics in common, such as their coagulation by treatment with heat, or by treatment with strong acids or alkalies. As the analytical methods improved, it was discovered that all these substances contained nitrogen, and this property soon was considered a distinctive characteristic of the group of compounds under study.

The first systematic investigation of these substances was undertaken in the third decade of the nineteenth century by Gerardus Johannes Mulder who examined such biological materials as silk, blood fibrin, egg white and gelatin. Early in his studies Mulder realized the significance of this group of nitrogen containing compounds stating:

> ". . .It is without a doubt the most important of the known components of living matter, and it would appear that, without it, life would not be possible." (Translation from German cited by Fruton and Simmonds, p. 15, 1959.)

The fundamental value of these materials led Mulder (1839) to give the name of "protein" (from the Greek proteios, of the first rank) to what he apparently believed to be a characteristic chemical radical (Vickery and Osborne, 1928). Later Liebig introduced such nonspecific terms as "albuminoids", "proteids", or "protein bodies" to describe these compounds. The generic term "proteins" as presently used to describe a particular class of substances, did not come into common usage until the first decade of the present century (Fruton and Simmonds, 1959).

Early in the nineteenth century, Francois Magendie (1816) demonstrated that the nitrogen present in the body was derived from the nitrogen compounds present in the food. Later, Magendie (1841) obtained the first evidence which indicated that not all proteins have the same nutritional value, demonstrating that gelatin could not take the place of meat in the diet. Boussingault (1836) suggested that the nutritive value of foods of plant origin depends upon their nitrogen content and postulated that the nutritional potency of a food is proportional to its nitrogen content. Unfortunately, this postulate dominated the thinking of the leading investigators of the time for the remainder of the century, and consequently the interesting question posed by Magendie in 1841 was not investigated again until 1897, when Rubner recognized that proteins of different origin did not have the same value in nutrition (Block and Mitchell, 1946).

By this time it had been clearly established that proteins were made up of smaller units, the amino acids, and as early as 1876, Escher had demonstrated that the nutritive value of gelatin, as determined in dogs, could be improved by the addition of the amino acid tyrosine; by 1905 it had been shown that tryptophan and cystine were also necessary for the satisfactory nutrition of dogs fed a basal diet of gelatin (Maynard, 1951). Two years later, Thomas B. Osborne (1907) suggests for the first time that the relative proportions of the amino acids in the protein may determine its nutritional value.

These preliminary findings were followed by a series of brilliant protein nutrition studies initiated by Karl Thomas and followed by

Thomas B. Osborne, Lafayette B. Mendel and H. H. Mitchell which opened new horizons for the investigation of the nutritional values of foods, laying the foundations for the modern concept of the role of proteins in nutrition (Thomas, 1909; Osborne and Mendel, 1917; Osborne, et al., 1919; Mitchell, 1923-24 a, b, c). Of special interest is the work of W. C. Rose which resulted in the classification of amino acids as essential or nonessential, according to whether they can or cannot be synthetized by the organism, leading subsequently to the establishment of quantitative requirements for the essential amino acids (Rose, 1937; 1938). All this work has led to the formulation of the modern theory of an amino acid balance or unbalance in the composition of a given protein and its effect upon the nutritive value of this protein (Elvehjem and Krehl, 1955; Flodin, 1957; Harper, 1958). The known requirements for different animal species, including man, are listed in Table 1.1, and the essential amino acid pattern of the FAO reference protein is given in Table 1.2.

1.2.2 Structure and Basic Properties. Like carbohydrates and fats, proteins contain carbon, hydrogen and oxygen; in addition, they contain nitrogen in a fairly constant percentage (15-18%). Most of the proteins also contain sulphur and a few contain phosphorus. These basic elements are combined to form complex molecules of different forms permitting a gross classification of the proteins into various groups. For the purpose of this review it is sufficient to indicate that proteins can occur as simple proteins, conjugated proteins or derived proteins. Simple proteins are made up of amino acids exclusively, while

Table 1.1. Approximate requirements of essential amino acids when all other amino acids are present in adequate amounts. Calculated as per cent of diet (Block, 1956)

Approximate protein level	Rat young	Rat adult	Pig young	Dog adult	Human adult	Chicken young	Chicken adult	Turkey young	Turkey adult
%	20	10	13	6.5	10?	20	15	24	28
Arginine	0.2	0?	0.3	0.4	0	1.2		1.4	1.6
Histidine	0.4	0.1	0.4	0.1	0?	0.3			
Lysine	1.0	0.2	1.1	0.5	0.6	0.9	0.5	1.3	1.5
Tryptophan[a]	0.2	0.1	0.2	0.1	0.2	0.2	0.15	0.22	0.26
P-Alanine[a]	0.9	0.2	0.7	0.6	0.9	0.9			
Cys + Met	0.8	0.3	0.8	0.5	0.9	0.5	0.3	0.8	0.9
Threonine	0.5	0.3	0.6	0.5	0.4	0.6			
Leucine	0.8	0.3	1.2	0.8	0.9	1.4	1.0		
Isoleucine	0.5	0.5	0.6	0.5	0.6	0.6	0.5	0.7	0.8
Valine	0.7	0.4	0.6	0.7	0.6	0.8			
Glycine[a]	0	0			0	1.5		0.9	1.0

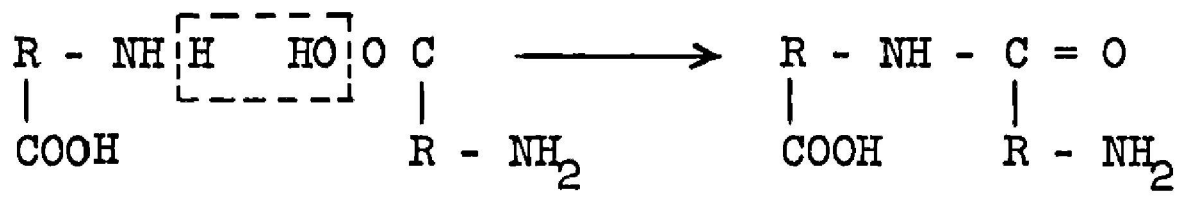
[a]Numerical values for tryptophan, phenylalanine, methionine and glycine may have to be modified due to failure of the investigators to take account of niacin, tyrosine, cystine and serine in the experimental diets.

Table 1.2. Essential amino acids in provisional pattern and milk and egg proteins (FAO, 1957)

Amino acids	Provisional pattern	Cow's milk	Human milk	Egg
	.. grams amino acid per 100 grams protein ..			
Isoleucine	4.2	6.4	6.4	6.8
Leucine	4.8	9.9	8.9	9.0
Lysine	4.2	7.8	6.3	6.3
Phenylalanine	2.8	4.9	4.6	6.0
Tyrosine	2.8	5.1	5.5	4.4
Sulfur containing:				
Total	4.2	3.3	4.3	5.4
Methionine	2.2	2.4	2.2	3.1
Threonine	2.8	4.6	4.6	5.0
Tryptophan	1.4	1.4	1.6	1.7
Valine	4.2	6.9	6.6	7.4
	.. miligrams amino acid per gram nitrogen ..			
Isoleucine	270	407	411	428
Leucine	306	630	572	565
Lysine	270	496	402	396
Phenylalanine	180	311	297	368
Tyrosine	180	323	355	274
Sulfur containing:				
Total	270	211	274	342
Methionine	144	154	140	196
Threonine	180	292	290	310
Tryptophan	90	90	106	106
Valine	270	440	420	460

conjugated proteins are combined with a nonprotein radical; derived proteins, as suggested by the name, include compounds which represent altered or degraded products of naturally occurring proteins through the action of enzymes, heat or chemical agents.

The amino acids, the basic units in the protein structure, are joined together through a special type of chemical bond, known as the peptide linkage (-NH-C=O), which occurs as illustrated below:



It is assumed that the proteins found in nature are built from their basic amino acids through a series of peptide linkages, but the exact process involved in their synthesis is not known (Flodin, 1957; Downes, 1955). In the case of plants and some of the lower forms, including yeasts and bacteria, nitrates and ammonium salts can be used as the raw materials for protein synthesis. In the case of most animals, however, the essential amino acids necessary for the synthesis of a given protein must be furnished preformed and be present simultaneously at the site of synthesis (Fruton and Simmonds, 1959). Ruminants are an exception because through their rumen bacteria they are able to utilize to some extent urea and nitrates in fulfilling their protein requirements (Fingerling, 1937; Reid, 1953). Nonessential amino acids can be synthesized in the body from other compounds present in the food or from products of intermediary metabolism (Rose, 1937; 1938).

Since the proteins contain free amino and free carboxyl groups they can combine chemically with either bases or acids. The tendencies

to acidic and basic dissociation are equalized at a particular hydrogen ion concentration, the isoelectric point, which is unique for each protein; at this point the protein can be most easily precipitated by salt solutions or by the addition of alcohol. Proteins in solution have colloidal properties, and in addition some authors believe that proteins can bind ions, both physically by adsorption and by uniting with them chemically. The solubility of proteins in aqueous solutions is variable, and they are all insoluble in the common fat solvents. Proteins can be precipitated from solution by a wide variety of substances. The salts of heavy metals and strong mineral acids are good precipitants, while fairly large quantities of various neutral salts are necessary to produce precipitation. The coagulation of proteins probably involves dehydration, and is brought about by enzymes (Downes, 1955; Fruton and Simmonds, 1959; Maynard, 1951).

The preceding remarks should make it clear that proteins are far from being stable and inert compounds, but on the contrary, they are quite labile. As a result, their nutritive value can easily be altered by physical or chemical agents. There is ample evidence that heat can enhance the nutritive value of certain proteins (Almquist and Merritt, 1952) while it may jeopardize the value of others (Murlin et al., 1938; Mitchell and Block, 1946). Finally, storage and processing may often result in a net loss in nutritive value (Mitchell et al., 1945).

1.2.3 General Metabolism. Ingested proteins pass to the stomach, where they are mixed with the appropriate enzymes, and hydrolysis into component amino acids is initiated. Although some absorption of amino

acids may take place at this level, most of it occurs in the small intestine, where the hydrolytic degradation of proteins into component amino acids is completed. Not all the nitrogen that is absorbed is in the form of individual amino acids, for some of the simpler peptides which result from incomplete hydrolysis, can be absorbed through the intestinal wall (Guyton, 1956; Crampton and Lloyd, 1959). Depending on digestibility factors, a residue containing variable amounts of undigested food proteins is passed in the feces. The feces contain also a certain portion of nitrogenous substances derived from secretions necessary for the digestive process. Additionally, there is in the feces some nitrogen contributed by the natural wear and tear of the intestinal walls, and a small amount contributed by the normal intestinal flora (Mitchell, 1923-24a). It should be clear that the nitrogen in the feces does not constitute just a residue of dietary nitrogen, but that it constitutes a mixture of dietary nitrogen and that which is generally called "metabolic" nitrogen. Schneider (1935) has pointed out that the metabolic nitrogen in the feces of rats and pigs can be divided in two fractions: a digestive fraction which varies directly with the amount of food dry matter consumed, and a constant fraction which is probably of true excretory origin. The latter fraction, at least in rats, seems to be related to body size, and more so to body surface; in a sense this fraction could be called the endogenous nitrogen of the feces. However, in the case of the human Schneider (1935) points out, there is no constant fraction in the fecal metabolic nitrogen. Some authors believe that the degradation of the enzyme proteins

and other secretions will tend to maintain an amino acid mixture of constant composition in the gastro intestinal tract regardless of the type of diet consumed, and consequently, it is assumed that the amino acid balance in the food is not critical for absorption (Block, 1956; Flodin, 1957).

The absorbed nitrogenous products, enter the organism via the portal circulation, pass to the liver where they participate actively in anabolic and catabolic reactions, undergoing oxidation, deamination, transamination, decarboxylation and all the typical reactions of the intermediary metabolism of proteins. In the case of mammals, including the human, the principal product of excretion as the terminal product of nitrogen metabolism is urea, but some waste nitrogen is also passed in the urine as ammonia, uric acid, creatinine, and free amino acids. Under pathological conditions other products may be present, or the relative amounts of one or more of the compounds listed may be excreted in abnormally high proportions (Fruton and Simmonds, 1959; Guyton, 1956).

Early in the study of protein metabolism, Folin (1905) noticed that there were certain end products (e.g., creatinine) that were excreted in the urine in fairly constant amounts, while others tended to be related to the amount of protein in the diet. This observation led him to formulate the theory of two forms of protein catabolism, a variable one which he called "exogenous catabolism" and a constant one which he denominated "endogenous catabolism". The two mechanisms were considered to be independent, and were defined separately. The "exogenous

catabolism" was thought to reflect the level of protain in the diet, and hence its variable nature. The "endogenous catabolism" was thought to reflect those catabolic processes which are the basic functions of a living cell. Accordingly, the latter was thought to be constant, and dependent on such factors as the size of the individual and the corresponding metabolic activity. This was interpreted to mean that once the proteins were incorporated into a tissue, they simply grew old and were then replaced, but could not be used for the synthesis of other tissues (Block, 1956).

With the advent of radioactive isotopes, it was possible to demonstrate the dynamic state of body proteins. Employing this technique, and using ℓ(-)-Leucine containing two isotopes, Schoenheimer, Ratner and Rittenberg (1939) found that less than one third of the absorbed nitrogen appeared in the urine, 57% had served for replacing body proteins, and a small amount was present in the non-protein fraction. Different body proteins revealed different activities with respect to their acceptance of the labeled leucine. Amino acid samples isolated from different body proteins indicated that different chemical processes take place continuously in the body proteins, for at least 32% of the dietary leucine had found its way to replacing 24% of the liver leucine and 7% of the carcass leucine. Isotopic nitrogen was found in all the amino acids with the exception only of lysine, and only one third of the N^{15} was deposited together with the carbon chain. All this evidence suggests an extensive and continuous shifting which occurs even under nitrogen balance conditions, and the existence of

what the authors called an "amino acid pool". Thus the sharp distinction between "endogenous" and "exogenous" nitrogen in the sense proposed by Folin (1905), is no longer valid, but it remains open to question if the dynamic state of the body proteins changes in any manner the end result of catabolism (Mitchell, 1955). As Mitchell (1943) points out,

> "there is nothing in Schoenheimer's work that denies the existence of a constant type of catabolism of nitrogen-containing compounds in the tissue."

The concentration of essential amino acids in the blood is surprisingly constant, even after trauma or periods of poor nutrition (Flodin, 1957). The amino acid pattern in the radial circulation, however, is quite different from the pattern in the portal vein, but just the same, there is good correlation between the blood amino acids and the amino acid content of the food. Amino acids are present in tissue cells at a greater concentration than in the extracellular plasma. The rate of accumulation within the cell depends on the rate of anabolic activity within the cell; the greater this activity within the cell, the faster the rate of accumulation of amino acids within the cell. These observations suggest active transport of amino acids involving a carrier of limited capacity, probably a pyridoxal derivative (Flodin, 1957). Because of the low capacity of the carrier, there is likely to be competition of amino acids, and Flodin (1957) suggests that it is at this metabolic level where the balance of amino acids seems to be critical.

Protein synthesis is a function of ribonucleic acid-rich cell fractions, such as the microsomes. Incorporation of amino acids requires

magnesium ions, ATP, an ATP precursor, and a buffer, but there is no competition among amino acids for incorporation into the body proteins. It has been shown that this incorporation involves true protein synthesis, and an enzymatic mechanism involving carboxyl activation has been proposed (Flodin, 1957).

A diagramatic summary of the salient features of protein metabolism discussed in this section is presented in Figure 1.1.

1.2.4 Proteins as Foods. The special place of proteins in the diet has been recognized for over a century, but many of the intricacies of their multiple functions in the organism remain to be clarified.

Proteins must be furnished by food. They are irreplaceable for many vital functions, and for the biological synthesis of these proteins to occur, in most animal species, all the necessary basic units, the amino acids, must be present simultaneously at the site of synthesis (Fruton and Simmonds, 1959). About one third of the amino acids of which proteins are composed are said to be essential (Table 1.1), that is, they must be supplied in the food, for they either cannot be synthesized by the organism, or if they can be synthesized, this does not occur at a sufficiently rapid rate to meet the metabolic demands of the living system (Rose, 1937; 1938).

Foods are often classified as "energy foods", "body building foods" or "protective foods". Proteins are principally "body building foods", but they may serve as "energy foods" if necessary, illustrating once again the importance of interaction among dietary components. It is

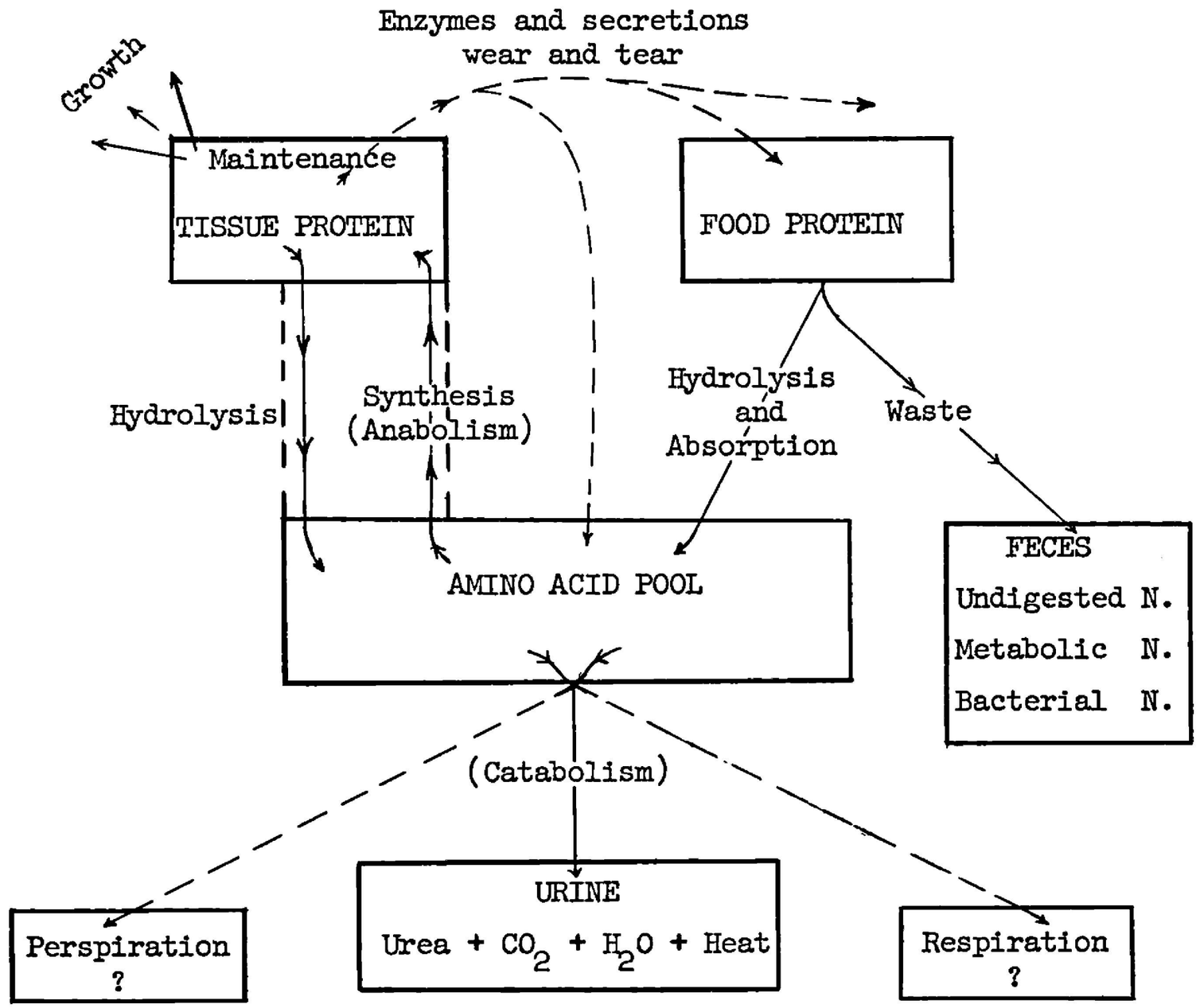


Figure 1.1. General metabolism of food protein

also becoming increasingly evident that proteins may have some "protective" functions. Studies on the relation of nutrition and infection indicate that nitrogen utilization may be markedly affected even prior to the onset of the clinical manifestations of common infections (Scrimshaw, et al., 1959; Gandra and Scrimshaw, 1961; Wilson, et al., 1961).

In general terms, however, the utilization of proteins as foods seems to follow common patterns among different species. Thus, the protein requirements for growth for the rat, chick and pig are similar, and the biological value of egg albumin, whole egg, beef muscle, wheat gluten, casein and peanut flour as determined in growing rats and mature humans, are also very similar (Mitchell, 1954; 1959). In this connection, the Committee on Amino Acids of the National Research Council of the United States (1959) offered the following comment:

> "Although different criteria have been used in both the experimental design and the evaluation of the data, two striking findings result: a) the proportions of amino acids required are similar in all species, and b) the requirements for the indispensable amino acids are surprisingly low. This uniformity in the pattern of amino acids required is interpreted to be a reflection of the similarities in the amino acid composition of the tissues of different species."

In order to achieve a better understanding of the role of proteins as foods, the methods and procedures employed in establishing protein values should be examined closely, and the nature and properties of the mathematical models used for estimation studied. In the next chapter the standard procedures employed in the assessment of protein values are discussed.

CHAPTER II

METHODS FOR THE NUTRITIONAL EVALUATION OF PROTEINS

2.1 Introduction

From the discussion in the preceding chapter, it should be evident that the primary function of proteins in the diet is to furnish the organism with amino acids for the synthesis of tissue proteins. The relative adequacy of a given protein in making the required amino acid pattern available to the organism will determine the magnitude of observable biological responses (Osborne and Mendel, 1917; Mitchell, 1923-24a; Allison, 1953). Accordingly, practically all the methods concerned with the evaluation of proteins as foods, are directly or indirectly concerned with the appraisal of the relative efficiency of different proteins in satisfying requirements (Osborne, et al., 1919; Mitchell, 1923-24a; Allison, 1955; 1957; Barnes and Bosshardt, 1946; Murlin, et al., 1938).

In general, most of the methods currently employed in the evaluation of proteins are concerned with the comparison of biological responses obtained under carefully controlled experimental conditions. There are, however, multiple choices of either single responses or combinations of these which can be studied. The usefulness of any particular method employed for the assessment of protein values depends rather critically on the success in insuring that under the experimental conditions employed, protein is the only limiting factor for the response(s) chosen for study.

2.2 Biological Evaluation of Proteins--Basic Procedures

In the biological assessment of protein values two basic procedures are commonly applied. The first to be considered in this section utilizes "growth"[1] as the observable response, while the second basic method makes use of nitrogen balances in metabolic trials.

2.2.1 Efficiency Ratios. In early attempts for establishing the nutritional value of different proteins, Osborne and Mendel (1917) used ordinary graphs of the course of growth with time. Under ordinary normal conditions, growth should occur in a balanced fashion and for the total organism it should be the sum of the growth of its parts (Thompson, 1942). In this sense it should be possible to relate any measure of partial growth to total growth since it is reasonable to assume that these would differ in amounts specified by appropriate proportionality constants. Accordingly, it is possible to utilize different criteria for estimating growth. Two criteria commonly employed for this purpose are the simple measurements of height and weight; of these, and particularly in animal experiments, weight is the most commonly employed single measurement for the estimation of growth (Maynard, 1951).

In the early experiments of Osborne and Mendel (1917), sequential weight measurements of the experimental animals were used for the construction of the graphs of the course of growth with time. In these experiments, the test animals (rats) were fed standardized diets which

---

[1] "Growth" is used in a general sense to refer to the different ways in which an organism may increase its size.

presumably differed only in their sources of protein; under the experimental conditions it was expected that the qualitative differences of the proteins under study would be reflected in the observed differences in total growth or in the patterns of growth as estimated through the weight measurements.

Although the weight method permitted a crude separation of proteins of widely differing qualities, it soon became apparent that it did not permit critical discrimination. The uniform and orderly fashion in which growth proceeds on an average basis, was not always attainable with the relatively small size of the experimental groups which had to be utilized. Additionally, other factors, such as the level at which the proteins were fed, seemed to influence the results, magnifying or minimizing the differences observed in the response chosen for study.

Based on the experience gained in the course of these experiments, Osborne, Mendel and Ferry (1919) proposed that a more critical and accurate discrimination of proteins could be possible if the response observed, weight gain in a period of time, were to be expressed on the basis of the protein consumed in that period of time. In proposing this approach, the authors showed that the variation apparent in the absolute weight gains was greatly reduced when expressed in the form suggested (Osborne, et al., 1919). The index proposed by Osborne and Mendel for the evaluation of the biological value of proteins, is currently known as the "Protein Efficiency Ratio" (PER), and constitutes an expression of the weight gain obtained per unit of protein fed. Clearly its application is not necessarily restricted to weight gain, and may be used

when response criteria other than weight are considered. Alternatively, the reciprocal of the proposed index, which constitutes an expression of the amount of protein consumed per unit change in response, may be employed. Symbolically these two quantities can be expressed as follows:

$$\text{PER} = \frac{W_g}{P_c} \quad , \quad \text{or} \quad (\text{PER})^{-1} = \frac{P_c}{W_g} \quad , \qquad (2.1)$$

where W_g is the observed response, and P_c is the total protein consumed in a fixed period of time.

In practice, the utilization and interpretation of protein efficiency ratios, or its reciprocals, should be conditioned by the following considerations:

1. The comparison of proteins for their growth promoting effect is real only when the concentration of the protein under investigation is the only variable.
2. The food intake is regulated in large part by the caloric requirements of the animal, and these in turn depend on body size. This implies that the food intake increases with growth and that animals growing at different rates consume unequal amounts of food.
3. Individual differences exist even under the same conditions. To minimize the effect of these, the animals used in testing should be balanced at the start of any trial for initial weight as well as any other characteristics which are known sources of variability.

4. Several experiments covering a range of protein concentrations are necessary to establish the maximum growth promoting value of a given protein.

The considerations listed are, according to Osborne, Mendel and Ferry (1919), a sine qua non for the application of protein efficiency ratios in the evaluation of proteins.

Although properly derived protein efficiency ratios should be useful at least for initial screening of proteins, the possibilities of this method have not been fully exploited in common practice. At present, protein efficiency ratios are determined under ad-libitum feeding regimens and using a single protein concentration, usually 10% of the diet. Food consumption is commonly measured on a group basis, and no estimate of experimental error is provided with the estimates of the protein efficiency. The length of the experimental period is variable and depends on the animal species used for the trials; using rats, however, these experiments usually last from 30 to 40 days. Using mice as the experimental animal, Bosshardt and associates (1946) have shown that the length of time required for these experiments can be reduced to 10 days without a serious loss of information, so that the method could be profitably employed as a rapid preliminary screening technique.

The measurement of food consumption and weight gain on an individual basis is not difficult, and when both the food consumption and the weight gain are measured on an individual basis, estimates of experimental error are available for each of the two components involved in the calculation of the protein efficiency ratio. From these estimates

4. Several experiments covering a range of protein concentrations are necessary to establish the maximum growth promoting value of a given protein.

The considerations listed are, according to Osborne, Mendel and Ferry (1919), a sine qua non for the application of protein efficiency ratios in the evaluation of proteins.

Although properly derived protein efficiency ratios should be useful at least for initial screening of proteins, the possibilities of this method have not been fully exploited in common practice. At present, protein efficiency ratios are determined under ad-libitum feeding regimens and using a single protein concentration, usually 10% of the diet. Food consumption is commonly measured on a group basis, and no estimate of experimental error is provided with the estimates of the protein efficiency. The length of the experimental period is variable and depends on the animal species used for the trials; using rats, however, these experiments usually last from 30 to 40 days. Using mice as the experimental animal, Bosshardt and associates (1946) have shown that the length of time required for these experiments can be reduced to 10 days without a serious loss of information, so that the method could be profitably employed as a rapid preliminary screening technique.

The measurement of food consumption and weight gain on an individual basis is not difficult, and when both the food consumption and the weight gain are measured on an individual basis, estimates of experimental error are available for each of the two components involved in the calculation of the protein efficiency ratio. From these estimates

a first order approximation of error for the efficiency ratio can be obtained by squaring the total differential of the functional form for the protein efficiency ratio given in equation (2.1) and taking expectations. Using this procedure, the variance for the protein efficiency ratio is given by:

$$V(PER) = \frac{1}{P_c^2} V(W_g) + \frac{W_g^2}{P_c^4} V(P_c) - \frac{2W_g}{P_c^3} Cov(W_g, P_c), \qquad (2.2)$$

which may be more conveniently written as:

$$V(Q) = \frac{1}{P_c^2} \left[\sigma^2_{W_g} + Q^2 \sigma^2_{P_c} - 2Q\rho\sigma_{W_g}\sigma_{P_c} \right], \qquad (2.3)$$

where Q is the estimate of the protein efficiency ratio and ρ is the correlation coefficient for weight gain and protein consumed. Using this estimate of variance, and assuming Q is normally distributed confidence intervals for the efficiency ratio can be found in the usual way:

$$Q \pm t_\alpha \sqrt{V(Q)} \quad , \qquad (2.4)$$

where t_α is the appropriate "student t" statistic.

An approximate expression for the variance of the reciprocal of the efficiency ratio can be obtained using the procedure outlined; the result is given by interchanging W_g for P_c and vice versa in equation (2.2). Confidence intervals for the reciprocal can then be found using the appropriate form of equation (2.4).

The procedure outlined for estimating variances is particularly useful for the researcher with experience in biological trials, for he can draw from previous experiments to stipulate reasonable guesses for the variance in weight gain and protein consumed under a given set of circumstances. From these values then, reasonable estimates for the variance of an efficiency ratio can be derived for different values of ρ, when as often happens, this is not provided with the results. This approach can be especially useful in the planning of new experiments. Tukey (n.d., 1958?; n.d., 1960?) has recently studied the properties and general reliability of this approximate procedure for the estimation of variances concluding that in most instances it gives satisfactory approximations for the purposes at hand.

When the protein consumption is fixed, the second and third terms in equation (2.2) vanish, and as would be expected, the estimate of variance for the efficiency ratio under this restriction is given by:

$$V(Q) = 1/P_c^2 \left[V(W_g) \right] \qquad (2.5)$$

In general, the correlation between the weight gains and the protein consumed can be expected to be positive within the boundaries commonly used in the actual experiments. When there is perfect correlation between the gains observed and the protein consumed, then all $W_g = QP_c$ and $\sigma^2_{W_g} = Q^2\sigma^2_{P_c}$; in this situation $V(Q) = 0$, as would be expected. The variance of Q is a maximum when there is no correlation between the gains and the protein consumed. In practice the correlation between gains and the protein consumed is generally high so that the variance of the efficiency ratio should be small and sensitive comparisons possible.

Although the discussion in this section is concerned specifically with the protein efficiency ratio as proposed by Osborne, Mendel and Ferry (1919), it will be seen later that it also applies in the case of other ratio estimates for the relative efficiency of proteins as foods.

2.2.2 Nitrogen Balance. The basic concept of balance studies in biology is no different from the idea of balance in other disciplines, and it is concerned with the establishment of net gains or losses in the utilization of the materials needed by the organism. Although reports in the literature date to the seventeenth century the first indications of the use of this method in biological research (Maynard, 1951), the first real biological balance experiment probably was carried out by Boussingault (1838) in the latter part of the first half of the nineteenth century. In nutrition this method has proved extremely valuable and has been extensively used for the study of the metabolic pathways and the establishment of requirements for many of the known nutrients (Monroe, 1949; Crampton and Lloyd, 1959). In this respect, the metabolism of protein, and consequently that of nitrogen, is no exception. Nitrogen is constantly lost from the body and taken in as food; the determination of nitrogen in the food and excreta should provide, therefore, a quantitative measure of protein metabolism indicating whether the organism loses or gains protein under a given set of experimental conditions.

By definition, then

$$B' = I - (F + U), \qquad (2.6)$$

where B' is the apparent nitrogen balance, I is the nitrogen intake,

F the nitrogen in the feces and U the urinary nitrogen. Although losses of nitrogen occur through pathways other than the feces and urine, these are generally ignored in nitrogen balance experiments because of the extreme difficulty to measure them. Mitchell and Hamilton (1949) have shown, however, that some nitrogen may be lost through perspiration in hot, humid environments. It is conceivable, therefore, that neglecting other pathways of nitrogen excretion may result, under certain conditions, in the misinterpretation of data.

The apparent nitrogen balance, B' in equation (2.6), is said to be positive whenever $I > (F + U)$, and in this case a net gain in nitrogen occurs. If $I = (F + U)$, then $B' = 0$ and there is neither a gain nor a loss in body nitrogen; in this case the individual is said to be in nitrogen equilibrium. On the other hand, if $I < (F + U)$, a nitrogen loss is occurring, B' is negative, and the individual is said to be in negative nitrogen balance (Thomas, 1909; Mitchell, 1923-24a, 1944; Allison, 1955; FAO, 1957).

Nitrogen balance is really the sum of gains and losses from all the tissues of the body, and it should be clear that it is possible to have a positive nitrogen balance with a loss of nitrogen occurring in some tissues; thus under conditions of stress, for example, although some tissues may be maintained at the expense of others, an over-all positive balance may still be possible (Allison, 1955). In parallel, maintenance of nitrogen equilibrium does not mean that every tissue is being maintained, just as the provision of sufficient nitrogen to maintain nitrogen equilibrium in the adult does not necessarily imply the

fulfillment of nitrogen requirements. Nitrogen equilibrium can shift through metabolic adaptation so that it is possible to maintain an individual in a depleted state (FAO, 1957).

Thomas (1909) proposed a method to utilize information derived from nitrogen balance studies for the assessment of the quality or value of proteins in nutrition. This approach was later studied further, improved, developed and extensively applied by Mitchell (1923-24 a, b, c; 1943; 1944; 1954; 1955; 1959). By this procedure it is possible to estimate two important parameters in the evaluation of protein quality: a) Digestibility (D), defined as the percentage of intake of nitrogen which is absorbed; and b) Biological Value (BV), the percentage of the absorbed nitrogen which is retained in the organism. Symbolically these two quantities are:

$$D = \frac{I - F}{I}$$

and (2.7)

$$BV = \frac{I - F - U}{I - F} \quad ,$$

where I, F and U stand as defined for equation(2.6). Note that these two expressions bear a close resemblance to the protein efficiency ratios discussed in the previous section. The response observed and the stimulus applied are different, but nevertheless, these ratios are estimates of the relative efficiency of proteins in promoting a specified biological response.

The symbolic expressions for the apparent nitrogen balance, digestibility and biological value as given in equations (2.6) and (2.7) can

be improved by taking into account the contributions of the metabolic and the endogenous nitrogen to the total nitrogen measured in the feces and in the urine respectively. Introducing this modification, equations (2.6) and (2.7) may be written as follows:

$$B = I - (F + U) + (F_k + U_k);\ D = \frac{I - (F - F_k)}{I},$$

and

$$BV = \frac{I - (F - F_k) - (U - U_k)}{I - (F - F_k)}, \qquad (2.8)$$

where B is the nitrogen balance, D the digestibility, BV the biological value, F the total fecal nitrogen, U the total urinary nitrogen, F_k the metabolic nitrogen in the feces, and U_k the nitrogen of endogenous origin in the urine. Notice that the product BV x D, expresses the nitrogen retention as a percent of the intake. This index, known as the Net Protein Value (NPV), is also useful in the assessment of protein quality (Mitchell, 1923-24a).

Theoretically all the quantities in equations (2.8) should be measured in a single trial and simultaneously. However, in practice this is not possible and the metabolic and endogenous portions of the nitrogen metabolism are measured in separate trials in which the experimental animals are fed nitrogen-free diets, and then applied to the test diets. In the method originally proposed by Thomas (1909), it is assumed that both of these quantities are characteristic constants unaffected by dietary factors. This assumption is only approximately true and holds only when certain conditions are satisfied in the

experimental procedure. Mitchell (1923-24a) has pointed out that the composition of food, and particularly the presence of undigested materials, may affect the estimate of metabolic nitrogen present in the feces. To compensate for this possible effect, he suggests the equalization of crude fiber content in the nitrogen-free and the test diets; additionally, the estimate of the metabolic nitrogen in the feces for the test diet should be calculated for the total food consumed on the basis of the nitrogen in the feces per unit of food consumed on a nitrogen-free diet. Similarly, it should be pointed out that the nitrogen of endogenous origin determined by indirect means in the urine, may be affected by the caloric content of the diet when the supply of calories is close to the maintenance requirement, but if the caloric content of the diet is enough above the maintenance requirement, there is no detectable effect of calories on the nitrogen of endogenous origin in the urine.

From the preceding discussion it is evident that the determination of nitrogen balance is a direct method for studying the degree of utilization of dietary proteins for those metabolic purposes for which they alone can serve. The determination of nitrogen balance can be a tedious and slow process, but nevertheless, it has been used constantly as a method of protein assessment, and animals as well as humans have been used as subjects, both under normal and pathological conditions (Mitchell, 1923-24a; Allison and Anderson, 1945; Allison, et al., 1946; Leverton, et al., 1956; Sénécal, 1958; Gómez, et al., 1958). In using this method, however, certain conditions must be satisfied:

1. The diet must contain only the protein or mixture of proteins under investigation, and should not contain non-protein nitrogen other than that present in the food under investigation.
2. The over-all composition of the diet must be so adjusted that no protein will be metabolized for energy, except insofar as such utilization is conditioned by the inability of the protein under investigation to maintain the nitrogen integrity of the tissues, failure to promote growth and/or the inability to satisfy the requirements for other specific physiologic functions.
3. A special effort should be made to avoid the intermixing of food with excreta, as well as the mixing of urine and feces.
4. There should be a short period of adaptation to the diet to be tested prior to the actual collection of information. The same subjects should not be used continuously through a long series of different experimental periods, especially when testing proteins of low value, to avoid the carry-over effects of adaptation to previous test diets.
5. The preparation of the food, especially when working with human subjects, should receive special attention, since the taste of the food and possible losses through vomiting are particularly important factors in this case.

In addition to the conditions listed, careful consideration should be given to the physiological status of the subjects (Mitchell, 1923-24a; Scrimshaw, 1961). Also, although there is some evidence that

neither the order (Mitchell and Carman, 1924) in which different foods are tested for their biological value, nor the relative frequency of the food offerings to the subject (Mitchell, 1923-24a; Chanutin and Mendel, 1922) have an effect upon nitrogen balance, recent nitrogen balance studies conducted in children recovering from kwashiorkor indicate that both the order in which different therapeutic diets are tested and the length of the time intervals between feedings may influence the results of nitrogen balance studies in human subjects (INCAP, 1961).

In practice each investigator satisfies these conditions in accordance with the facilities at his disposal, for there is no uniform set of rules to be followed. Thus, for example, the balance periods used in evaluating proteins are variable both in length and number; it is customary to employ one two or three three-day balance periods, but some investigators prefer to employ other combinations of duration and number of balance periods (Bressani, et al., 1958, 1960; Wissler, et al., 1948). Such differences in experimental procedure complicate the interpretation and comparison of data obtained in different laboratories. Additionally, as was pointed out in the case of the protein efficiency ratio, no estimates of variance are provided with the results of nitrogen balance studies. Most of the time these results are presented in the form of graphs constructed with average figures. In the case of nitrogen balance studies in humans, individual results are customarily presented, and their interpretation is based primarily on the consistency of directional changes on repeated trials; the magnitude of these

changes is of secondary importance because it is felt that they are greatly influenced by large individual differences.

Using the same approximate procedure suggested in the case of the protein efficiency ratio, it is possible to find expressions for estimating the variance for the calculated indices of digestibility and biological value. Expressions for the variance of the nitrogen balance and the nitrogen absorbed present no special problem and can be written directly, for both of these quantities are simple linear forms of I, F and U. The metabolic and endogenous nitrogen fractions in the feces and the urine do not contribute to the variance of these quantities for they are only suitable correction constants. The variance of the nitrogen balance is given by:

$$V(B) = \sigma_I^2 + \sigma_F^2 + \sigma_U^2 - 2\left[\rho_{IF}\sigma_I\sigma_F + \rho_{IU}\sigma_I\sigma_U + \rho_{FU}\sigma_F\sigma_U\right], \qquad (2.9)$$

and the variance of the nitrogen absorbed is:

$$V(A) = \sigma_I^2 + \sigma_F^2 - 2\rho_{IF}\sigma_I\sigma_F \,. \qquad (2.10)$$

In practice, however, the application of these two equations is conditioned by the lack of information concerning the nature and behavior under different conditions of the various correlation factors needed. Measures of balance and absorption are obtained individually in most experiments, and direct estimates of the variance of these quantities are then possible. These estimates can be utilized to derive estimated approximate expressions for the variances of digestibility and the biological value by using the following relations:

$$V(D) = \frac{1}{I^2}\left[\sigma_A^2 + D^2\sigma_I^2 - 2D\,\rho\,\sigma_A\sigma_I\right] \qquad (2.11)$$

and,

$$V(BV) = \frac{1}{A^2}\left[\sigma_B^2 + \overline{BV}^2\sigma_A^2 - 2\overline{BV}\,\rho\,\sigma_B\sigma_A\right] \qquad (2.12)$$

In all these equations all symbols stand as previously defined.

The correlation of nitrogen intake and digestibility and that of biological value and absorption can be expected to be positive and relatively high, so that these two expressions may be employed in practice if reasonable values for the correlation factors can be assumed within the suggested range. Under these conditions, the expressions given might be particularly useful in exploratory work for the planning of new experiments.

2.3 Biological Evaluation of Proteins--Modifications of Basic Procedures

Modifications in the basic procedures used for evaluating protein quality pertaining to both methodology and interpretation have been suggested from time to time. Some of these modifications will be considered in this section.

2.3.1 Replacement Values. Murlin, et al., (1938) suggested that if nitrogen balance is interpreted using a reference protein of good quality such as egg or milk, the time necessary for conducting the actual trials could be reduced eliminating the standardization period with low protein diets. The results of these experiments are given as "replacement values" (V_R), calculated from data obtained while feeding sufficient nitrogen to maintain equilibrium or positive balance according

to the following expression:

$$V_R = 100 - \left[\frac{B_r - B_t}{I_r}\right](100) \qquad (2.13)$$

where V_R is the replacement value of the protein tested, B_r the nitrogen balance for the reference protein, B_t the nitrogen balance for the test protein and I_r the nitrogen intake for the reference protein.

2.3.2 Depletion and Repletion. It is reasonable to assume that the regeneration of tissue in the adult takes place in a similar fashion as the growth of new tissue in the young animal. If an adult subject is sufficiently depleted of its protein stores, the rate of tissue regeneration can be rapid and approach the rate at which growth of new tissue proceeds in the young. Under these premises Cannon and associates (1944), Wissler, et al. (1948) and others (Vivanco, 1960; Bressani, 1960), have determined the nutritive value of proteins by first depleting and then repleting the protein stores of the adult rat. Depletion is accomplished by feeding a protein-free diet until the rats have lost 25% of their initial body weight. The animals are then fed the test diets and the rate of repletion measured; seven days repletion periods are usually sufficient for the estimation of nutritive value. There is an excellent correlation between weight gain in the repletion period and the regeneration of blood, liver, and carcass proteins, so that weight recovery alone can be a measure of nutritive value (Allison, 1955). Weight gains, weight gains expressed as efficiency ratios, and nitrogen balance have been widely employed as the response in the

evaluation of protein quality by this method with satisfactory results (Frazier, et al., 1947; Bressani, 1960; Vivanco, 1960).

2.3.3 Protein Minima. Melnick and Cowgill (1937) proposed that the determination of the minimum amount of nitrogen or protein necessary to maintain nitrogen equilibrium could be used as an index of the nutritive value of dietary proteins. These authors determined this minimum by plotting observed nitrogen balance results against nitrogen intake, or nitrogen absorbed, in the region of nitrogen equilibrium and interpolating to the point of nitrogen intake or absorption which results in equilibrium or zero balance. In proposing this procedure, Melnick and Cowgill (1937) recognized that many factors can influence the estimate of the protein minima, emphasizing the importance of standardizing the test subjects to similar steady states when conducting such trials for the purpose of comparing one protein minimum with another.

2.3.4 Nitrogen Balance Indices. Allison and Anderson (1945) have demonstrated that the relation between nitrogen balance and nitrogen absorbed is in fact curvilinear, but the locus of the observations for absorption in the region of negative balances and low positive balances is sufficiently well approximated by a straight line. This relationship can be written as:

$$B' = K\,(A) - NE_o\,, \qquad (2.14)$$

where B' is the apparent nitrogen balance as defined in equation (2.6), A is the nitrogen absorbed and NE_o is the nitrogen excretion on a protein-free diet; the regression coefficient (K) is closely related to the

biological value (BV) of Mitchell. Allison and Anderson (1945) call this regression coefficient the "nitrogen balance index", and point out that it is identical with biological value (BV), when the nitrogen excretion on a protein-free diet (NE_o), is actually the sum of the metabolic and the endogenous nitrogen (F_k and U_k respectively) previously discussed. The regression coefficient (K) is a constant for a given protein when the experimental observations are obtained within the limits where the linear relationship for nitrogen absorbed and nitrogen balance can be expected to hold. Additionally, it is reasonable to assume that nitrogen balance is a reflection of protein quality, and hence the constant K may be used as an estimate of the protein quality. One advantage in this approach is that it follows simple regression theory, and accordingly estimates of the constants and their variances are immediately available, and confidence intervals and tests of hypothesis present no real problem.

Allison and associates (1946) have shown that a linear relationship also holds for nitrogen intake (I) and nitrogen balance in the region of negative or low positive balances. This relationship can be written as:

$$B' = K' (I) - NE_o \quad , \qquad (2.15)$$

where B', as before, is the apparent nitrogen balance, I is the nitrogen intake and NE_o is the nitrogen excretion on a protein-free diet. The regression coefficient (K') in this case is the equivalent of the net protein value (NPV), when NE_o is truly the sum of the metabolic and endogenous nitrogen. In this situation the ratio $(\frac{K}{K'})$ is the

digestibility factor (D), defined in equation (2.8). Algebraically, these relations follow after subtracting equation (2.15) from equation (2.14) to obtain:

$$K(A) - K'(I) = 0$$

and (2.16)

$$\frac{K'}{K} = \frac{A}{I} = D \ ,$$

where all symbols stand as previously defined. Alternatively this relation can be obtained by substituting $(\frac{A}{D})$ for (I) in equation (2.15). Also note that since:

$$NPV = (BV)\ D \ ,$$

and (2.17)

$$(BV) = K \ ,$$

then,

$$NPV = KD \ ,$$

and (2.18)

$$K' = KD = NPV \ .$$

2.3.5 Net Protein Utilization. More recently, a rapid method for the determination of the net protein value (NPV) has been developed by Miller and Bender (1955). Since these authors utilize a different approach than the one suggested by Mitchell (1923-24a), the index obtained using the procedure of Miller and Bender (1955) is known as the "net protein utilization" (NPU). Clearly, NPU is equivalent to NPV but the calculated values of the two indices are not necessarily numerically

equal. In the method of Miller and Bender (1955), nitrogen balance is estimated by carcass analysis rather than of the urine and feces; rats are used as the experimental animal. When the assays are carried out under standardized conditions with respect to the level of protein fed and amounts of fat, minerals and vitamins in the diet, the term "net protein utilization" (standardized), or NPU_{st}, is applied to the results of these assays (Platt and Miller, 1959). Under certain conditions, there is a very high correlation ($r = 0.936$) between carcass water and carcass nitrogen content (Braham, et al., 1959). These authors have obtained satisfactory results using chicks and predicting the carcass nitrogen from carcass water using a linear relationship, thus avoiding the tedious task of digesting the complete animals for carcass nitrogen determinations.

2.3.6 Other Modifications. Other proposed versions of the basic procedures for the evaluation of protein quality consist principally in utilizing different response criteria. Changes in enzyme systems, in serum protein fractions, in liver solids, and in the rate of fat deposition in the liver, for example, have been used in the past with relatively good success. Similarly, different animal species, including bacteria and protozoa, have been used as test subjects to assess the nutritive value of proteins (Allison, 1955). For the purpose of this dissertation, however, the methods that have been briefly discussed should suffice. They illustrate that all the biological procedures for determining the nutritive value or proteins are concerned with measuring their relative efficiency in producing some observable response involving the conversion of dietary protein to tissue protein.

2.4 Non-biological Evaluation of Proteins

In the preceding sections it has been pointed out that the nutritive value of a protein or a mixture of proteins depends largely on the quantity and proportion in which these provide the essential amino acids to the organism. Knowing the requirements for a given species in terms of proportions and amounts of amino acids it should be possible, then, to estimate the nutritive value of a protein on the basis of its amino acid composition. In this section some of the procedures that have been proposed for utilizing information on amino acid content in determining protein values will be presented.

2.4.1 Method of Kühnau. Kühnau (1949) proposed the use of human milk as a reference for the evaluation of the nutritive value of dietary proteins and suggested the following scheme for utilizing available information in terms of essential amino acid composition. Let R_i (i=1,2,3,...10) be the content per 16 grams of nitrogen of the i^{th} essential amino acid in human milk, the protein used as reference; let T_i (i=1,2,3...10) be the concentration per 16 grams of nitrogen for the corresponding i^{th} essential amino acid in the test protein, then an estimate of nutritive value is given by:

$$\frac{\sum_{i=1}^{10} T_i}{\sum_{i=1}^{10} R_i} = TV \,, \qquad (2.19)$$

where TV is defined as the "total value" (gesamte Wertigkeit) of the test protein relative to human milk. Kühnau later noted that his proposed "total value" credits the test protein with concentrations of

amino acids in excess of those in the standard or reference protein, excess which presumably contributes nothing to the nutritive value of the test protein. Accordingly he suggested that the "total value", calculated according to his original scheme, should be corrected as follows. Let t_j ($j=1,2,\ldots n$; $n \leq 10$) be the excess amino acid concentration in all cases where $R_i < T_i$, then a corrected estimate of nutritive value can be obtained by

$$\frac{\sum_{i=1}^{10} T_i - \sum_{j=1}^{n} t_j}{\sum_{i=1}^{10} R_i} = PV , \qquad (2.20)$$

where PV is defined as the "pure value" (reine Wertigkeit) of the test protein relative to human milk. Additionally Kühnau defines the difference between the "total value " (TV) and the "pure value" (PV) as the supplementary value (Erganzungswertigkeit), which is presumably a measure of the availability for supplementary relations with other proteins which may be deficient in the amino acids occurring in excess in the test protein.

2.4.2 Method of Oser. A method similar to that of Kühnau (1949) has been proposed by Oser (1951). In this case whole egg protein instead of human milk is used as a reference standard. The nutritive value of a test protein is estimated by calculating what Oser calls the "essential amino acid index" (EAAI) according to the following expression:

$$100 \sqrt[10]{\prod_{i=1}^{10} \frac{T_i}{R_i}} = \text{EAAI} , \qquad (2.21)$$

where T_i is the concentration of the i^{th} essential amino acid in the test protein, and R_i is the concentration of the corresponding essential amino acid in the whole egg protein. In this case as before, excess amino acid concentrations in the test protein are disregarded by arbitrarily defining as one those ratios in which $R_i < T_i$. The use of ratios and the geometric mean should result in a better distribution of weights for the contribution of each amino acid to the calculated index.

2.4.3 Method of Mitchell and Block. Mitchell and Block (1946) have proposed a method which approaches the subject of the chemical evaluation of proteins in a different manner, and does not use the concept of integrated essential amino acid content applied by Kühnau (1949) and Oser (1951). Using egg protein as a standard, Mitchell and Block (1946) compared the percentage of each essential amino acid in the test protein to that of the reference standard they had chosen. The largest relative deficit with respect to the standard was taken as the "chemical protein score", since this should be the amino acid limiting the utilization of the total protein. These percentage deficits are highly correlated with the corresponding biological values determined by the nitrogen balance method ($r = -0.861$) while there is little or no correlation of the chemical scores with the digestibility of the proteins studied.

Later the chemical score was redefined as the complement of the original definition given above, and Mitchell (1954) reported that a

considerable improvement in the correlation with biological value (r = +0.948) was possible by using a modified version of the method of Oser (1951). The modification consisted in calculating the geometric mean of the relative deficits (or of their complements) of all the essential amino acids present in the test protein in smaller relative quantities than in egg protein. The high correlations suggest that the nutritive value of a protein is determined essentially by its amino acid composition; furthermore, the improvement in the correlation observed when using the modified essential amino acid index would seem to suggest that the single amino acid most limiting in the protein does not determine the nutritive value of the protein in toto, although it is probably responsible for the greater part of it.

2.4.4 Method of "Protein Scores" (FAO). A new concept concerning the protein used as a standard in the assessment of protein values has come about in recent years with the introduction of the FAO reference protein (FAO, 1957). Actually, the reference protein is an artificial pattern of essential amino acids (see Table 1.2) elaborated by the FAO Committee on Protein Requirements, on the basis of the present knowledge of the role and requirements of amino acids in nutrition. In proposing this amino acid pattern as a reference protein, the Committee made the following comment:

> "There is no experimental evidence that the provisional pattern, based on minimal requirements as these are now known, is superior to the patterns found in good food proteins such as those of milk and egg. A case could be put forward for adopting one or other of the latter as standards of comparison in assessing protein quality. The Committee has, however, preferred to adopt the former." (FAO, 1957).

Associated with this pattern, the FAO Committee proposed the calculation of a protein score, along the lines of reasoning followed by Mitchell (1954) in proposing the use of the chemical score. The FAO protein score, which reflects protein quality, is taken as the ratio of the lowest amino acid in the protein under consideration to the corresponding amino acid in the proposed provisional amino acid pattern. The protein score is conceptually the complement of the chemical score as originally defined by Mitchell and Block (1946) and equivalent to it as redefined; it differs from these in using a different standard for comparison.

CHAPTER III

THEORETICAL DEVELOPMENT

3.1 Basic Premises

3.1.1 Equivalence of Different Procedures. Some of the methods commonly used in the assessment of the nutritive value of proteins were briefly discussed in the previous chapter. It was pointed out that these methods differ primarily in the type of response chosen for study and in the experimental procedure employed. Ultimately, however, they all yield a numerical quantity which is used as an index of the nutritive value of different proteins, and which, independently of the procedure followed, often results in a similar ranking order for a given set of proteins. This implies that the indices of protein quality estimated for the same protein by different procedures are roughly proportionally related. This also implies that functional relationships exist among different response criteria. Knowledge of these functional relations would permit the calculation of the necessary proportionality constants for the interconversion among the different types of indices employed in the assessment of protein quality.

The estimates of protein efficiency, however, are affected by the level of dietary protein used in the feeding trials, and thus, the comparison of efficiency indices for different proteins, estimated at different levels of intake, does not constitute a fair comparison of their quality. Some information on the quality of proteins can be obtained from the comparison of efficiency indices estimated at the same level of intake. It should be clear, however, that such information on

protein quality applies only at the level of dietary protein used to obtain the efficiency estimates, and cannot be generalized over all levels of intake. For a generalized interpretation of protein quality using efficiency indices, it becomes necessary to study the comparative behavior of the proteins under consideration in a range of suitably chosen levels of dietary protein. Osborne, Mendel and Ferry (1919) proposed that this approach should be followed in applying their scheme for the biological evaluation of proteins in order to determine the points of maximum efficiencies; these authors further indicated that the comparison of the biological efficiencies of different proteins was valid only when made between maximum efficiencies.

Although maximum efficiencies are an important consideration in protein quality, maximum efficiencies per se are not sufficient for the characterization of protein quality. Allison (1954) has suggested that the shape of the efficiency curves should be of value for the identification and characterization of the different factors that affect protein quality. This approach is particularly appealing for it would make possible the development of a generalized theory which would be applicable to the different types of responses generally employed in the estimation of the biological efficiency of proteins. The systematic testing of the hypothetical model with existing information on protein quality and with results from experiments designed for the investigation of the properties of the model, should permit the introduction of theoretical refinements in the original hypothesis making possible the formulation of an improved model which may reflect more accurately the modus operandi of the biological system under consideration.

3.1.2 The Meaning of Protein Quality. The biologically determined quality of a given protein[2] is the result of the direct effect and interaction of many factors. Not all the factors involved in the measurement of the quality of a protein are inherent to the protein itself, and dietary as well as non-dietary factors may be reflected in the estimates of protein quality. Thus, the physiological state of the test individual, which may or may not be related to nutrition, may condition his response. Moreover, the relative adequacy of the diet with respect to non-protein nutritive entities, can also affect the utilization of the protein. Clearly, under such conditions, the responses obtained cannot be considered good measures of the quality or nutritive value of protein per se. For this reason, the adequate control of all the known non-protein factors which may influence the estimates of the nutritive value of proteins should be a primary consideration in the experimental procedures followed in the course of feeding trials.

Under proper experimental conditions the quality of the protein is determined by factors inherent to the protein itself. This, however, does not guarantee that unique estimates of nutritive value for a given protein will be obtained in repeated trials. As it was pointed out in Chapter I, proteins are labile and can be very sensitive to physical and environmental factors. Thus, prolonged storage and/or differences in

---

[2] In the discussion which follows, the term protein should be understood to refer to individual as well as to mixtures of proteins from different sources unless otherwise specified in the text.

processing prior to the test for quality, for example, can result in real differences in nutritive value.

It is generally agreed that digestibility is important for the evaluation of protein quality, but there is no general agreement as to whether this factor should be included or not as an inseparable part of the numerical values used as indices of protein quality. Thus, digestibility is not explicitly considered in the protein efficiency ratio estimated according to the scheme of Osborne, Mendel and Ferry (1919), but it is explicitly considered in the estimate of biological value proposed by Mitchell (1923-24a). In the latter case digestibility is estimated separately and used as a conditioning factor for the biological value in the calculation of the so-called "net protein value". Such discrepancies in concept, however, should not hinder the proper interpretation of data on protein quality, provided digestibility is recognized as a pertinent factor, and the necessary information given so that it can be taken in consideration when necessary. In practice, the digestibility of many of the materials tested is high and confined to a fairly narrow interval, so that its net contribution to the indices of quality is relatively small in most cases.

The quality of a protein is a function of its amino acid pattern and content. The magnitude of the responses commonly used in the assessment of protein quality must vary, therefore, in accord with the assortment and proportions of amino acids present in the protein and made available to the organism. The more complete the assortment of amino acids, and the closer the proportions approach the optimum in terms of requirements, the higher the quality of the protein.

To date, twenty-six amino acids have been identified as products of the hydrolysis of different proteins. Although all of the amino acids are in some sense essential within the organism, not all of them are needed in the diet; as was pointed out in Chapter I, certain amino acids can be synthesized in the course of metabolism from other amino acids and, in certain species, from other nitrogen sources. The term "nonessential" used to describe these amino acids refers to the fact that it is not strictly necessary that they be furnished as preformed entities by the dietary protein. On the other hand, depending on the species, about ten amino acids which cannot be synthesized, or cannot be synthesized at a sufficiently rapid rate to satisfy metabolic requirements, must be furnished as preformed entities by the dietary protein; these are ordinarily called the "essential" amino acids.

The spectrum of essential amino acids may vary with changes in the basic physiology of the organism. For example, the amino acids that are considered essential for the growing individual are not necessarily the same amino acids considered essential for the maintenance of the general well being of an adult individual. Both of these sets of essential amino acids may in turn differ from the set of amino acids considered essential under stress conditions (such as disease or trauma), when some metabolic pathways may be partially or totally blocked, or extensive tissue repair may be necessary. In every case, however, the amino acids are ultimately used for the synthesis of proteins for growth, repair, maintenance of the <u>status quo</u>, or any combination of these processes.

Since growth is the physiological state which probably imposes the greatest demands for amino acids, it is not unreasonable to assume that a protein of good quality for growth is also satisfactory under other non-pathologic physiological conditions. Accordingly, growth will be used as the response of choice in attempting to describe and illustrate (responsewise) the relation of the amino acid pattern and content to the quality of proteins.

Assuming that the nonessential amino acids are available in adequade amounts, and considering the essential amino acids as a group, without identification of the individual acids in the group, the multiplicity of possible situations which may arise in practice could be reduced to the following general classes:

A. An ideal protein which contains all the essential amino acids in the correct proportions[3] in optimum concentrations.

B. A group of proteins which deviates from the ideal only because of suboptimal amino acid concentrations.

C. A group of proteins which deviates from the ideal because of an unbalance in the proportions of their amino acids.

D. A group of proteins which lack one or more of the essential amino acids.

Evidently, such a scheme of classification is neither the only one possible nor is it pretended to be exhaustive in the consideration of

---

[3]Proportion refers to the relative ratios among the essential amino acids in a protein which independently of absolute quantities establish the so-called "amino acid pattern" for that protein.

all the factors which may affect the quality of proteins. However, it should be useful as a basis for discussion and for illustrating the complexity of the functional system under which the quality of proteins is established.

The ideal protein in class A would be expected to produce the largest responses at optimum rates, resulting in maximum efficiencies at minimum dietary protein concentrations. Theoretically, the proteins included in class B should produce results similar to those obtained with proteins in class A, because in this case it should be possible to correct the quantitative deficiencies in amino acids by increasing the protein content in the diet. Maximum efficiencies can be expected to be lower and to occur in this case at higher dietary protein concentrations than in the case of the ideal protein.

The expected responses in the case of the proteins included in class C will vary according to the type of unbalance present in the protein. If the unbalance is not marked enough to result in competition among the amino acids, the response obtained with these proteins should be equivalent to that obtained with proteins included in class B. When the amino acid unbalance is marked enough to result in competition among the amino acids, the responses will be smaller and as the protein concentration in the diet increases should approach a lower attainable maximum response than the proteins in class B. The maximum efficiencies for these proteins can be expected to be lower and to occur at higher dietary protein concentrations than for the proteins included in class B. Toxic effects of amino acids are known to exist (Harper, 1958), and

some types of unbalance may induce such effects. In this case, an inhibition of response would occur at the higher levels of dietary protein. However, since relatively large concentrations of amino acids are likely to be necessary to produce toxic effects, this type of unbalance will not be considered in this thesis. The proteins included in class D may be regarded as a special case of the proteins in class C, since the lack of an essential amino acid can be regarded as a limiting type of unbalance. Feeding proteins of the type included in class D would result in losses instead of gains.

The content of nonessential amino acids may be an important conditioning factor to the quality of proteins which has not received adequate attention in the past. It is not unreasonable to expect that a nonessential amino acid deficit is capable of altering the nature of the responses described above for different classes of proteins. In general, a deficit of nonessential amino acids can be expected to be detrimental to the quality of balanced proteins, while it may be either detrimental or beneficial to the quality of unbalanced proteins. The effect on the latter would depend on the course of selective utilization of essential amino acids to supply the missing nonessential amino acids. Additional research is necessary to clarify the role of the nonessential amino acids in determining protein quality. The results from such studies may well justify the incorporation of the group of nonessential amino acids <u>in toto</u> as an additional unit in the frame of essential amino acids.

Although in the preceding discussion reference has been made to "unbalanced proteins", a deliberate effort was made throughout to avoid the use of the term "imbalance". It was felt that this concept has not yet been consistently and satisfactorily defined, and not infrequently is misused and applied unrestrictedly in situations where conditional responses are observed. Alternatively, the proteins themselves are considered as unbalanced, and the effects of unbalance on response are regarded ultimately as the result of amino acid deficiencies. It should be noted that in accord with the discussion in the previous paragraphs, these deficiencies may be either single or multiple, original, induced or conditional.

Further development of the concepts outlined in this section, coupled with the study of the nature of response curves over a broad range of dietary concentrations for different proteins, may give the best lead for the better understanding of the role of amino acids in the nutritive value of proteins.

3.2 Development of the Model

3.2.1 Characterization of Stimulus (Input) and Response (Output).

From the discussion in Chapter II, where the different procedures commonly used for the evaluation of proteins were presented, it should be evident that in the general case these techniques are concerned with the evaluation of measurable differences observed in different types of biological phenomena, when certain characteristics of the diet with respect to protein are changed.

At the beginning of this chapter it was pointed out that differences in procedure, and hence in the choice of biological change

measured in the trials as well as in the quantitative expression of the protein attributes of the experimental diets, should not result in real differences in the estimates of protein quality. Accordingly, in the discussion which follows, the different types of changes in biological phenomena which may be observed will be considered in a general sense as "responses" (output), while the various forms of specifying the protein attributes in the experimental diet will be considered as quantitative expressions of "stimuli" (input). Specific reference to a particular kind of response and stimulus will be avoided in as far as possible, but the general concepts should be applicable in all cases. Additionally, it will be assumed that the physiological status of the test individuals is satisfactory for the response to occur and all the essential nutritive entities in the diet, other than protein, are present in sufficient amounts to satisfy requirements. In other words, the experimental conditions are such that protein is the only factor limiting the biological response.

In the general case then, for a protein of given quality and in a fixed period of time, it should be possible to obtain responses which increase in magnitude as the intensity of stimulus increases. The increments in response obtained for corresponding increments in stimulus should be constant, or nearly constant, at the lower intensities, but they should gradually decrease in magnitude when the increments in stimulus occur in the range of higher intensities. Thus, the net responses should approach a maximum value rapidly at the lower levels of stimulus, and slowly at the higher levels of stimulus in a fashion

which suggests a biological behavior conditioned by a law of diminishing returns. This is not an uncommon situation in biological phenomena, and there is abundant experimental evidence which illustrates this point in general (Willcox, 1959) and in the case of proteins and amino acids in particular (Almquist, 1953; Bressani and Mertz, 1958).

Unfortunately, most of the experiments conducted for the purpose of the evaluation of protein quality, have been primarily concerned with the evaluation of proteins at the economically feasible levels for the feeding of domestic animals, and consequently are usually carried out using single levels of input, and do not permit the study of the nature of the response curves. The same situation prevails when humans are used as test subjects, for in this case it is practically impossible to study responses in a wide range of stimulus intensities. As a result, the data derived under such experimental conditions as described above, although valuable for the particular purposes of each trial, cannot be utilized in a broader sense for the characterization of the relevant parameters which may permit the general classification of proteins on the basis of the characteristics which determine their quality or nutritive value. Consequently, any classification scheme, such as the one used for discussion in Section 3.1.2, will have to be developed on hypothetical grounds and modified in accordance with the results of experiments performed to test the theory.

A general summary of the preceding discussion is presented graphically in Figure 3.1. The hypothetical response curves shown illustrate the general nature of the stimulus-response (input-output)

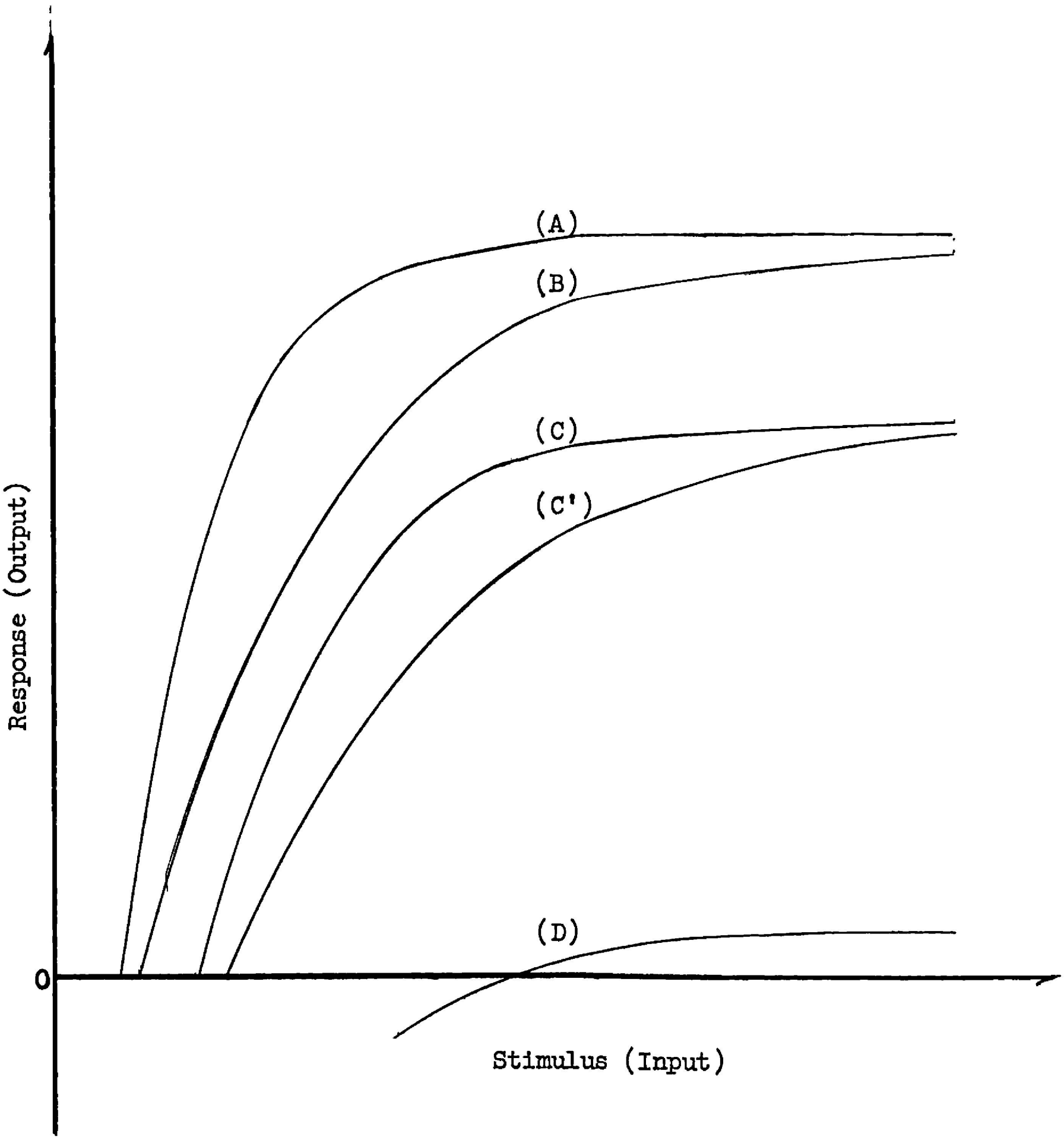


Figure 3.1. Theoretical responses for proteins of different quality. (A) Expected response for an "ideal" protein. (B) Expected response for a balanced protein with suboptimal concentration of amino acids. (C) and (C') Expected response for proteins with a competitive type of amino acid unbalance. (D) Expected response for an unbalanced protein approaching the limiting case (lack of one or more essential amino acids).

relations which might be expected in accord with the theoretical considerations outlined in Section 3.1.2.

3.2.2 Stimulus, Response and Relative Efficiency. The discussion in the previous section suggests that the quality or nutritive value of proteins may be determined satisfactorily through the characterization of the response curves obtained when feeding different proteins over a suitable range of concentrations under controlled experimental conditions. The determination of the nutritive value of proteins by this procedure, however, does not take into consideration the important economic aspects of protein nutrition. It would be desirable that some measure of the relative "cost" of the proteins be included as an integral part of the different indices of quality that may be calculated. This objective is usually accomplished in practice, for as it was pointed out in Chapter II, most of the procedures employed for the biological evaluation of the nutritive value of proteins result in the calculation of some kind of relative efficiency index which expresses the units of response attained on a per unit of input basis. Since cost can be assumed to be proportional to input, from such indices, it should be possible to make inferences concerning the best way to utilize different proteins to obtain maximum returns at minimum cost. The calculation of an efficiency ratio, however, does not eliminate the need for considering responses over a graded range of protein concentrations in the diet. Clearly, the efficiencies are not independent of the intensity of input, and consequently for the proper interpretation of results it becomes necessary to study the nature of the efficiency curves.

In the general case, the numerical value of the relative efficiency will increase as the input increases in the region where the ratio of the resulting increment in response for a corresponding increment in input is greater than the relative efficiency; there will be no change in efficiency when the ratio of the two increments is equal to the relative efficiency, and this will decrease when the ratio of the resulting increment in response to the corresponding increment in input is less than the relative efficiency. In other words, relative efficiencies will increase in value, reach a maximum and decrease in value as the input increases.

Under comparable experimental conditions proteins of good quality should reach a relatively high maximum efficiency at a fairly low level of input, while proteins of intermediate quality should generally result in lower maximum efficiencies attained at variable levels of input. In all cases, the efficiency maxima for proteins of intermediate and poor quality should occur at higher levels of input than the maxima for the proteins of good quality.

The beginning of positive responses in the continuum observed, i.e., positive relative efficiencies, should occur at lower levels of input stimulus for the proteins of good quality than for other proteins. The terminal relative efficiencies obtained at the higher levels of input are likely to vary according to protein quality.

A graphic summary of the preceding discussion is presented in Figure 3.2. The concepts considered are schematically illustrated using hypothetical relative efficiency curves.

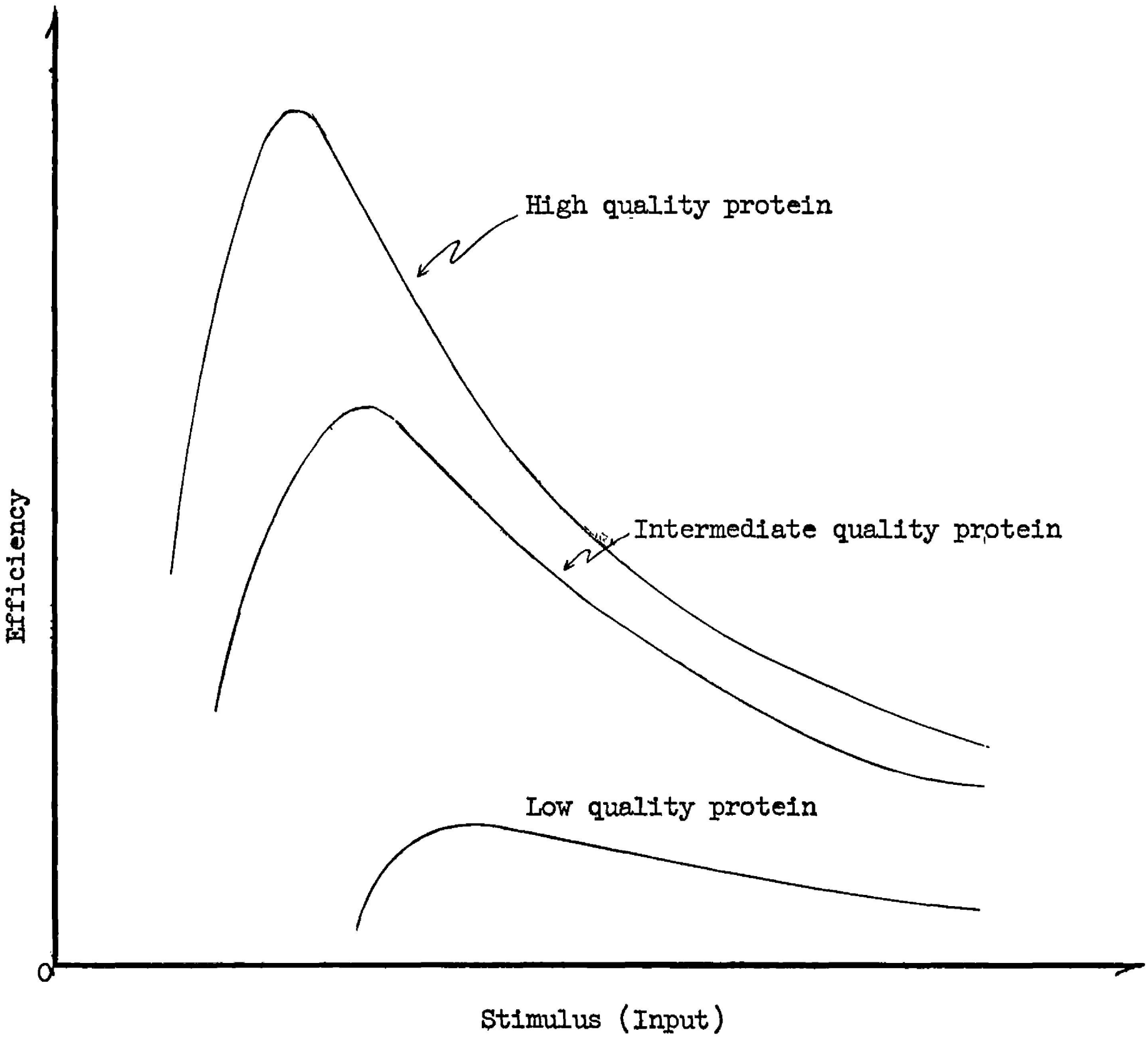


Figure 3.2. Schematic expected behavior of the relative efficiency curves for proteins of different quality.

3.3 Mathematical Derivation of the Model

3.3.1 Relative Efficiencies. Let Q be the biological response for a given species observed in a fixed period of time t, and let I be the cumulative input over the time t necessary to produce Q. When both Q and I are specified quantitatively using the same units of measurement, the relative efficiency, z say, for a given protein is given by

$$z = \frac{Q}{I} , \qquad -k_1 \leq Q \leq k_2 ; \qquad 0 \leq I \leq k_3 . \tag{3.1}$$

For a given species, the values of the limiting constants for Q, k_1 and k_2, depend respectively on the physiologic tolerance for net losses and the maximum biologic capacity for gain in the criterion used to estimate Q. The quantitative measure of the input, I, will take only positive values and will be limited by the food intake capacity, k_3, of the species, and the concentration of protein per unit of food.

Since it is not biologically tenable under the usual experimental conditions that I = Q = 0, there is only one admissible discontinuity in z; this occurs when $I = \bar{0}$. With proteins of very poor quality, Q may take only negative values, for these proteins are likely to result in losses only; other proteins should produce both gains and losses. The latter, which involve negative values of Q, should occur at reasonably low levels of input, and should become more marked as I decreases in value. Accordingly

$$\lim_{I \to 0} z = \lim_{I \to 0} \left[Q/I \right] = -\infty . \tag{3.2}$$

Increments in input ΔI, will produce corresponding increments in response ΔQ. When i (i=1,2,...n) levels of input are considered, the following conditions should hold:

$$z_{i+1} > z_i \quad \text{if} \quad \frac{\Delta Q_i}{\Delta I_i} > z_i \ , \tag{3.3}$$

$$z_{i+1} = z_i \quad \text{if} \quad \frac{\Delta Q_i}{\Delta I_i} = z_i \ , \tag{3.4}$$

and

$$z_{i+1} < z_i \quad \text{if} \quad \frac{\Delta Q_i}{\Delta I_i} < z_i \ . \tag{3.5}$$

These conditions follow immediately after writing the relations

$$z_i = \frac{Q_i}{I_i} \ , \quad \text{and} \tag{3.6}$$

$$z_{i+1} = z_i + \Delta z_i = \frac{Q_i + \Delta Q_i}{I_i + \Delta I_i} \ ,$$

solving for Δz_i and rearranging to obtain

$$z_{i+1} - z_i = \Delta z_i = \frac{\frac{\Delta Q_i}{\Delta I_i} - \frac{Q_i}{I_i}}{\frac{I_i}{\Delta I_i} + 1} \ . \tag{3.7}$$

Because of the physiological limitations inherent in any biological system, Q cannot be expected to increase indefinitely but will rather approach a maximum value which will vary depending on the species and the conditioning effect that the quality of the protein may have on the test subject. Accordingly, ΔQ can be expected to be relatively large

at the low levels of I and to become increasingly smaller as the biological maximum of Q is approached. This general nature of biological responses was illustrated in Figure 3.1. Since ΔI can be maintained fairly constant in the biological system, then $\frac{\Delta Q}{\Delta I}$ should satisfy (3.3) at the lower levels of I, then in a different range of I, as I increases, will satisfy (3.4), and finally at the higher values of I will satisfy (3.5). Therefore, z can be expected to increase in value, reach a maximum and then gradually decrease as I increases; the general nature of such curves was shown in Figure 3.2. It is customary to express the scale of input in these graphs as percent of protein in the diet, although the values of z are usually calculated on the basis of the absolute amounts of protein consumed or absorbed according as to whether the investigator wishes to estimate digestibility separately or not. The percent concentrations of protein in the diet constitute in a way an expression of input intensity which may be considered as common to all proteins. Since the scale change mentioned above involves only a simple linear transformation, it does not affect the general nature of the relationships discussed.

The equation for the calculation of relative efficiencies given in (3.1) can be alternatively written as a product as follows:

$$z = \left[\frac{1}{I}\right](Q) , \qquad (3.8)$$

emphasizing the two component nature of z. In addition, if the two components, $\left[\frac{1}{I}\right]$ and Q, can be expressed as functions of the protein concentration of the diet, p say, it should be possible to derive a

meaningful mathematical model for the efficiency of proteins of the general form

$$z = f_1(p) \cdot f_2(p) , \qquad 0 \leq p \leq 1 . \tag{3.9}$$

First approximations for these functional forms, which may be useful for a better understanding of the various factors involved in the determination of the nutritive value of proteins, will be derived in the next two sections.

3.3.2 Response (Output) as a Function of Protein Concentration in the Diet. Consider a response Q, which may be any suitable expression of a biological phenomenon, obtained in a fixed period of time t. It is assumed that the dietary protein, conditioned by the physiological capacity of the test individual, is the only limiting factor for the response Q. Let η be the average of the response maxima attained by all the individuals in a given species under specified experimental conditions. From previous discussion, it seems reasonable to assume additionally that any change in response produced by increasing p, the protein concentration in the diet, would be proportional to the response yet to be made in an amount γ determined by certain qualitative aspects of the protein under test, which pertain to the rate at which the maximum response is approached. In other words, the rate of change for Q with respect to p is given by

$$\frac{dQ}{dp} = \gamma(\eta - Q) , \tag{3.10}$$

or

$$\frac{dQ}{dp} + \gamma Q = \gamma \eta , \tag{3.11}$$

which is a simple first order differential equation of the general form

$$\frac{dy}{dx} + Ty = M \ . \tag{3.12}$$

The general solution for differential equations of the form in (3.12) is given by

$$y = e^{-\int TdT}\left[\int M\, e^{\int TdT}\, dT + C\right] . \tag{3.13}$$

Referring the expression in (3.13) to (3.11), the solution for Q in terms of γ and η is given by

$$Q = e^{-\int \gamma dp}\left[\int \gamma\eta e^{\int \gamma dp}\, dp + \xi^{*}\right], \tag{3.14}$$

which upon integration yields

$$Q = \eta + \xi^{*} e^{-\gamma p} \ , \tag{3.15}$$

where ξ^{*} is the constant of integration.

From previous discussion, it should be evident that Q will take negative values for dietary concentrations of a protein less than some value of p, p_b say. The value of p_b will vary for different proteins depending on quality, but it should be fairly low for proteins of good and intermediate quality. At the concentration p_b, neither gain nor loss will be evident in the biological phenomenon under consideration and Q will be equal to zero. The value of p_b then can be considered as the absolute minimum protein concentration necessary to prevent losses in any response studied. For values of p greater than p_b, gains will be

evident in terms of the response under consideration, and accordingly, Q will take positive values. To satisfy the preceding conditions, the following two restrictions are necessary for ξ^*:

$$\text{i)} \quad \xi^* < 0 \tag{3.16}$$

and

$$\text{ii)} \quad |\xi^*| > \eta \ .$$

Since ξ^* will take negative values only, equation (3.15), for the sake of convenience will be written in the form

$$Q = \eta - \xi\, e^{-\gamma p} \ , \tag{3.17}$$

where $\xi = -\xi^* > \eta > 0$. Equation 3.17 is the functional form $f_2(p)$ sought for Q. The graph of $f_2(p)$ for different values of p takes the general forms presented schematically in Figure 3.1.

3.3.3 Stimulus (Input) as a Function of Protein Concentration in the Diet. It has been stated previously that the stimulus (input) may be specified quantitatively in different ways. However, all the procedures employed ultimately will be expressions of the amounts of protein made available to the test subjects. Whether these amounts are expressed in terms of protein consumed, protein absorbed, or similar expressions in terms of nitrogen, should not influence the general nature of $f_1(p)$, since all these expressions are equivalent under simple linear transformations.

To characterize the stimulus in terms of the protein concentration in the diet, it is not unreasonable to assume, as a first approximation, that the rate of change of the stimulus with respect to the protein present in the diet will remain constant. Symbolically, then

$$\frac{dI}{dp} = \beta \ , \tag{3.18}$$

which on integration yields

$$I = \beta p + \alpha \ , \qquad 0 \leq p \leq 1 \ , \tag{3.19}$$

and the functional form $f_1(p)$ sought for the stimulus I, is

$$f_1(p) = \frac{1}{I} = \frac{1}{\alpha + \beta p} \ , \qquad 0 \leq p \leq 1 \ . \tag{3.20}$$

The biological interpretation of the parameters α and β in equation (3.19) will change depending on the particular choice in the quantitative measure of protein input (stimulus) into the system. In general, however, β will reflect some measure of the average food consumed, while α should be a general correction term for the amount of protein specified by the product βp. Further discussion on the nature of these parameters will be postponed, since a better understanding of the nature of α and β will be possible when the complete model is considered.

Equation (3.19) represents a straight line, and its reciprocal $f_1(p)$ as given in equation (3.20) is a decreasing monotone which will vary from $\frac{1}{\alpha}$ when $p = 0$, to $\frac{1}{\alpha + \beta}$ when $p = 1$.

3.3.4 The Complete Model. The two expressions, $f_1(p)$ as given in equation (3.20), and $f_2(p)$ as given in equation (3.17), can be combined according to equation (3.9) to obtain a mathematical model for the efficiency z, as a function of p, the protein concentration in the diet. Symbolically, the model is given by

$$z = \frac{\eta - \xi e^{-\gamma p}}{\alpha + \beta p} \quad , \qquad 0 \leq p \leq 1 \; . \tag{3.21}$$

The meaning of the structural parameters in the system will now be considered.

It will be recalled that η was defined as the average maximum attainable response for all the individuals of a given species under specified experimental conditions for a period of time t. Additionally, in considering on a theoretical basis the different types of responses which can be expected when feeding proteins of different qualities (section 3.2.1), it was pointed out that balanced proteins could be expected to produce the same response maxima, although these would be attained at different levels of input; unbalanced proteins on the other hand, should result in lower response maxima. Accordingly, η should be useful for identifying non-toxic types of unbalanced proteins, when the results obtained with different proteins can be referred to the value of η obtained with a protein of known good quality under the same experimental conditions. Egg and milk proteins have been used in the past for comparisons in the case of single point quality estimates, and should also be adequate for comparing different values of η when used as indicated above.

In deriving equation (3.17), γ was defined as the proportion of the gain yet to be made achieved by an increment in stimulus (input). It was also pointed out that the magnitude of response (output) increments produced by corresponding increments in stimulus (input) will gradually decrease as the response approaches its limiting maximum η;

the parameter γ is a measure of the rate of decrease of the increments in response, and consequently an indicator of the rate (with respect to p) at which the maximum response is attained. This is an important consideration in the question of protein quality which may permit the establishment of a quality gradient for proteins capable of producing the same maximum response. The integration constant, ξ, is a scaling factor which varies with η and γ.

Under ideal experimental conditions, all the digestible protein contained in the food should be utilized for the production of the response, after the relatively small protein needs for basic metabolism are satisfied. Additionally, the observed response should depend exclusively on the dietary protein, and should not be influenced by other factors in the diet. In practice, however, the ideal experimental conditions are seldom, if ever, fulfilled. The parameter α in the model measures, in protein equivalents, departures from ideal experimental conditions.

The parameter β measures the average "effective" food consumption during the experimental period, and is proportional to the observed food consumption. The proportionality constant which relates these two quantities, and which may be considered as a food efficiency factor, is a function of α. It may be derived as follows: Let F_c be the observed food consumption and let P_c be the protein consumed in the course of a given trial (all other symbols stand as defined previously). Obviously,

$$P_c = F_c p \tag{3.22}$$

and since P_c is equivalent to the protein input defined in equation (3.19) it follows that

$$\alpha + \beta p = F_c p \ , \tag{3.23}$$

and

$$\beta = F_c - \frac{\alpha}{p} \ . \tag{3.24}$$

Using the relation in equation (3.22), equation (3.24) may be rewritten in the form

$$\beta = F_c\left(1 - \frac{\alpha}{P_c}\right) \ , \tag{3.25}$$

where the portion in parenthesis is the proportionality constant relating the observed food consumption to the "effective" food consumption. Note the following relations which follow immediately from equation (3.25):

$$\begin{aligned} &\text{if} \quad \alpha < 0, \quad \text{then} \quad F_c < \beta \ , \\ &\text{if} \quad \alpha = 0, \quad \text{then} \quad F_c = \beta \ , \\ \text{and} \quad &\text{if} \quad \alpha > 0, \quad \text{then} \quad F_c > \beta \ . \end{aligned} \tag{3.26}$$

The preceding discussion suggests that additional biological research is necessary to be able to establish clearly the significance of all the parameters in the model. The estimates of the parameters obtained in the course of testing the model should give useful information in relation to the type of experiments that are needed for this purpose.

3.4 Behavior of the Model

The proposed model as given in equation (3.21) represents a family of curves which take the same general form of those presented in Figure 3.2. To illustrate further the behavior of the model, a family of efficiency curves, which may well belong to a group of proteins which allow the test subjects to attain the same maximum response but at different rates with respect to p is illustrated in Figure 3.3. These curves were obtained by taking fixed values for η, ξ, α, and β within realistic biological limits, and allowing γ to take different values; for each value of γ considered, the value of z was calculated for increasing values of p in a range from 0.04 to 0.40 at intervals of one-half of a one percent. Note the increasing sharpness of the maximum efficiencies which occurs simultaneously with a displacement to the lower values of p and the occurrence of higher maximum values. In general, these curves would be displaced to the higher maxima with increasing values of η when all the other parameters are fixed, while the maxima would be depressed with increasing values of β when all the other parameters are fixed.

For the sake of convenience, the discussion which follows will be limited to the range of p which produces positive responses. At the point of equilibrium in the system, when the stimulus applied results in neither gains nor losses, the value of z should be zero. In this case,

$$z = 0 = \frac{\eta - \xi e^{-\gamma p}}{\alpha + \beta p} , \qquad (3.27)$$

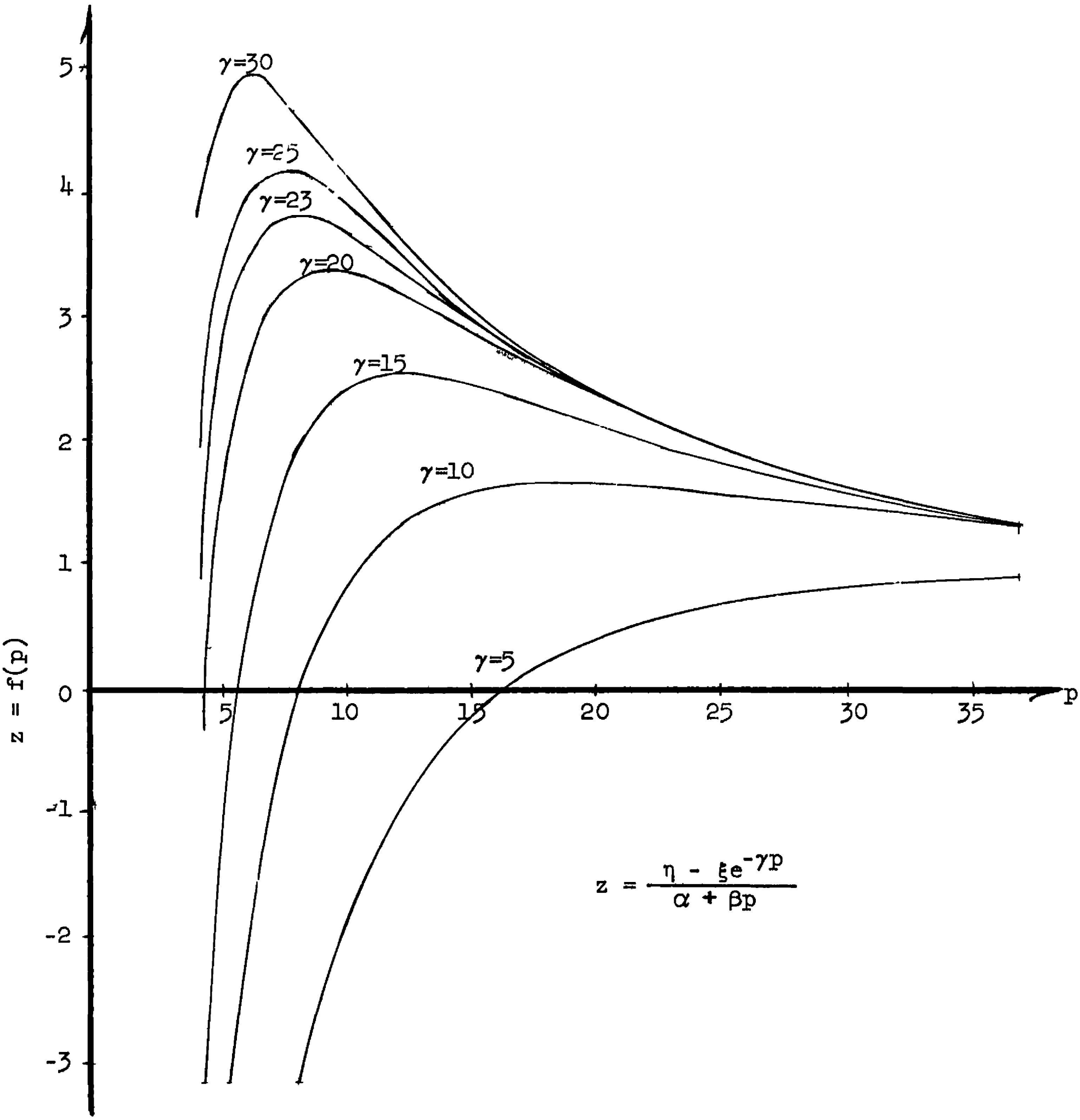


Figure 3.3. Behavior of $z = f(p)$ for different values of γ when η, ξ, α, and β are fixed.

and
$$\eta - \xi e^{-\gamma p} = 0 . \tag{3.28}$$

Then
$$\eta = \xi e^{-\gamma p} , \tag{3.29}$$
$$\ln \left[\frac{\xi}{\eta}\right] = \gamma p ,$$

and
$$p_b = \frac{1}{\gamma} \ln\left[\frac{\xi}{\eta}\right] . \tag{3.30}$$

When p is smaller than p_b (equation 3.30), the response will be negative, for it will occur in terms of losses. A state of equilibrium or gains in terms of response will occur in the system when p takes values that are equal to or greater than p_b.

Additionally, since in practice the concentration of protein in the diet seldom is higher than forty percent, it seems reasonable to consider this value as a practical upper limit for p. Hence, the discussion of the behavior of the model will be restricted to a range of values of p within

$$\frac{1}{\gamma} \ln\left[\frac{\xi}{\eta}\right] \leq p \leq 0.40 . \tag{3.31}$$

There are no discontinuities in z = f(p) as given in equation (3.21) within the restricted range of values specified in (3.31) which considersonly positive responses. It should be recalled that for the state of equilibrium in the system, or for positive responses to occur, it is necessary that

$$\left[\alpha + \beta p\right] > 0 . \tag{3.32}$$

The minimum value of z occurs at the lower limit in the range of p, and from equations (3.27) and (3.30) z minimum is obviously zero. At

the upper limit of p, z will take a finite positive value given by

$$z = \frac{\eta - \xi\, e^{-0.40\gamma}}{\alpha + 0.40\,\beta} \quad . \tag{3.33}$$

In general, this value will not be a maximum for z, since in most cases at p = 0.40 the response, $\eta - \xi\, e^{-\gamma p}$, should have reached already practically constant levels, and accordingly since the value of $\alpha + \beta p$ increases with p, z should be decreasing in value. Some exceptions will occur in the case of proteins which result in very low values of γ (see Figure 3.3); in such cases, z may still be increasing in value at p = 0.40.

In most cases, then, there will be an intermediate value of p within the limits under consideration which should result in a maximum efficiency. A solution for this value of p can be obtained by solving

$$f'(p) = 0 \ , \tag{3.34}$$

where f'(p) is the first derivative of z = f(p) with respect to p, and represents the rate of change in z per unit change in p.

From equation (3.21),

$$f'(p) = \frac{(\alpha + \beta p)\,\gamma\xi\, e^{-\gamma p} - (\eta - \xi\, e^{-\gamma p})\,\beta}{(\alpha + \beta p)^2} \ , \tag{3.35}$$

and equating to zero

$$\left[\alpha + \beta p_o\right] \gamma\xi\, e^{-\gamma p_o} = \left[\eta - \xi\, e^{-\gamma p_o}\right] \beta \ , \tag{3.36}$$

which can be rewritten in the form

$$\gamma\xi\ e^{-\gamma p_o} = Z_m\beta\ , \tag{3.37}$$

or

$$Z_m = \frac{\gamma\xi}{\beta}\ e^{-\gamma p_o}\ . \tag{3.38}$$

Equation (3.21) in the region of p considered has at most one inflection point, and since the minimum value of z is zero by definition, and the quantity on the right of equation (3.38) is positive and greater than zero, Z_m must be a maximum for z; the corresponding dietary protein concentration for z_m will be called the "optimum" dietary protein concentration, p_o. The equality in equation (3.38) cannot be satisfied in all cases, for as it was pointed out earlier, some proteins of low quality may not be capable of producing a maximum efficiency within the limits of p under consideration. It should be noted that Z_m occurs only with a single dietary protein concentration p_o, while other values of z, when z is not a maximum, may occur with two different values of p.

After some simple algebraic manipulations, equation (3.38) can be also written in the form

$$p_o = \frac{1}{\gamma}\ \ln\left[Z_m^{-1}\ \frac{\gamma\xi}{\beta}\right]\ . \tag{3.39}$$

Since p_o must be a positive quantity, it follows that

$$Z_m^{-1}\ \frac{\gamma\xi}{\beta}\ \geq\ 1\ , \tag{3.40}$$

and

$$Z_m \leq \frac{\gamma\xi}{\beta}\ , \tag{3.41}$$

which gives an absolute upper bound for the value of Z_m. Comparing the result given in equation (3.41) with the result given in equation (3.38) it is evident that the equality in (3.41) will hold only when p_o is equal to zero, which falls outside of the limits for the system under consideration. Therefore,

$$Z_m < \frac{\gamma\xi}{\beta} . \tag{3.42}$$

On occasions when a maximum efficiency is known to exist, explicit solutions for p_o and Z_m may be desired for the comparison of these values among different proteins. Such solutions can be obtained from equations (3.38) or (3.39), but would involve some rather tedious algebra. A simpler procedure for obtaining an explicit solution for p_o, and hence for Z_m since this value can be calculated once p_o is known using equation (3.21), can be outlined as follows. The numerator in equation (3.35) can be written in the form

$$\eta\beta e^{-\gamma p}\left[\frac{\xi(\beta + \alpha\gamma)}{\beta\eta} + \frac{\gamma\xi}{\eta} p - e^{\gamma p}\right] , \tag{3.43}$$

and the expression inside the brackets in equation (3.43) can be re-written as

$$(\alpha' + \beta' p) - e^{\gamma p} = \phi_1(p) - \phi_2(p) . \tag{3.44}$$

The solution for p_o is given when

$$\phi_1(p) - \phi_2(p) = 0. \tag{3.45}$$

The function $\phi_1(p)$ represents a straight line, while the function $\phi_2(p)$ is a simple exponential in p. A graph of $\phi_1(p)$ can be easily obtained by calculating the value of $\phi_1(p)$ for two suitably spaced values of p and joining them with a straight line, while the graph of $\phi_2(p)$ can be readily drawn using any standard table of values for the function e^x. When $\phi_1(p)$ and $\phi_2(p)$ are drawn using the same p scale for the abscissas, the intersection of the two lines gives the point at which the equality in equation (3.45) holds, and hence the solution for p_o; this can be read directly from the p scale in the graph. The schematic diagram presented in Figure 3.4 illustrates the general nature of the procedure.

The theoretical concepts developed in this chapter suggest that the proposed model allows a fairly good characterization of the efficiency curves used in the evaluation of protein quality. A systematic study of this model should permit, therefore, the beginnings of a better understanding of at least some of the critical factors which determine protein quality. A procedure for estimating the basic parameters in the model and their variances will be described in the next chapter.

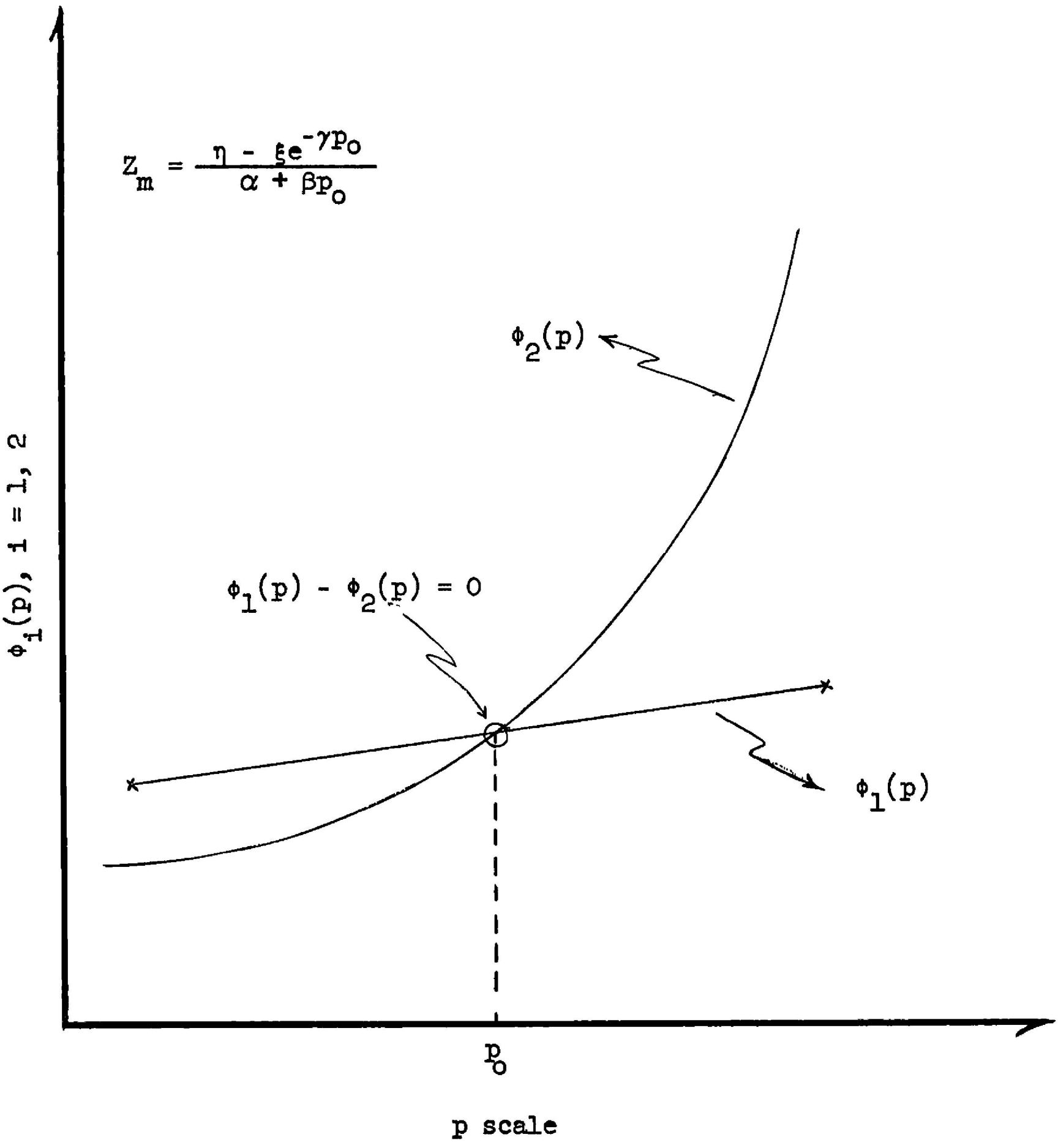


Figure 3.4. Schematic diagram illustrating a graphic method for the solution of p_o.

CHAPTER IV

ESTIMATION OF THE PARAMETERS IN THE MODEL AND VARIANCE OF THE ESTIMATES

4.1 Introduction

The data obtained in the course of biological trials for the assessing of the nutritive value of proteins consist of quantitative measurements of some biological response observed over a fixed period of time when the proteins under study are fed at different concentrations in the diet. Additionally, quantitative measurements reflecting the net intake over the period of study are also available. The choice of criteria and units used in the quantification of the data may vary, but it should be possible in all cases to put this information in a suitable form for calculating expressions of relative efficiency as defined in equation (3.1) in Chapter III. From the relative efficiencies and the dietary concentrations at which these were observed, estimates can be obtained for the parameters included in the model as given in equation (3.21) in the previous chapter. It should be noted, however, that because of the nature of the model, any parameter, with the exception of γ, may be divided out of the model without changing the basic structure of the system. It is evident, therefore, that all the parameters are not estimable, and at least one restriction is necessary to make estimation possible. In most cases this can be easily accomplished by specifying the value for one of the parameters in the model on the basis of previous knowledge, or by selecting some reasonable reference point for one of the parameters which may be chosen in a consistent manner in all cases.

A simple procedure for the estimation of the parameters and their variances in non-linear models will be presented in this chapter.

4.2 Estimation Procedure

4.2.1 The Principle of Maximum Likelihood. A procedure which uses the principle of maximum likelihood due to R. A. Fisher (1922), yields estimators of the parameters with many desirable properties. This procedure will be illustrated after introducing a new element in the model under consideration.

In the previous chapter a mathematical model for the relative efficiency of different proteins was developed and the biological significance of its parameters discussed. For the discussion in this chapter, it will be assumed that the errors in the model are normally and independently distributed with expectation zero and a constant variance σ^2. Thus, the model becomes

$$z = f(p) + \epsilon \,, \tag{4.1}$$

where f(p) stands as defined in equation (3.21) and ϵ is the residual random element mentioned above.

According to R. A. Fisher's definition of likelihood, the likelihood function for a series of N observations obeying equation (4.1) is given by

$$e^{L} = \left[\frac{1}{2\pi\sigma^2}\right]^{\frac{N}{2}} \exp\left\{-\frac{1}{2\sigma^2}\,\Sigma\left[z - \frac{\eta - \xi e^{-\gamma p}}{\alpha+\beta p}\right]^2\right\}, \tag{4.2}$$

where the summation is over the N observations in the sample.

Maximizing the likelihood is equivalent to maximizing the natural logarithm of the likelihood. This is given by

$$L = -\frac{N}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\Sigma\left[z - \frac{\eta-\xi e^{-\gamma p}}{\alpha+\beta p}\right]^2 . \qquad (4.3)$$

The maximum of L is obtained by differentiating equation (4.3) with respect to each of the unknown parameters, setting each partial derivative equal to zero and solving the set of simultaneous equations thus obtained.

The resulting set of equations will be non-linear in the parameters and, as it commonly happens in these cases, will not have explicit solutions for the values of the parameters which make the likelihood a maximum. There are several alternate general procedures which may be employed for solving such equations (Scarborough, 1950). Of these, a method due to Newton and first applied by Gauss for maximizing the likelihood embodies many desirable features. The procedure uses derivatives for successive approximations to the maximum point, and actually leads to least squares estimates of the parameters. However, under the usual assumption of normally and independently distributed errors with zero expectation and variance σ^2, the least squares estimates are the same estimates obtained by the method of maximum likelihood. This approach for obtaining maximum likelihood estimates for the non-linear case is commonly known as the Gauss-Newton method, although sometimes it is also called the Newton-Raphson method. The general procedure followed will be outlined in the next section.

4.2.2 The Gauss-Newton Method of Estimation. Consider a non-linear model of the general form

$$y = f(\underline{\theta}, x) + \epsilon \ , \tag{4.4}$$

where y is a dependent variable, $\underline{\theta}$ is a vector of r parameters to be estimated, x is an independent variable and ϵ is a random error NID $(0, \sigma^2)$. Substituting the maximum likelihood estimates $\underline{b}$ for the parameters $\underline{\theta}$, equation (4.4) becomes

$$y = f(\underline{b}, x) + e \tag{4.5}$$

where e is an estimate of the random element in the model. An estimate of an observation is given by

$$\hat{y} = f(\underline{b}, x) \ . \tag{4.6}$$

By Taylor's theorem any continuous function such as the one given in equation (4.6) can be expanded in terms of its partial derivative as follows:

$$\hat{y} = y_o + \left.\frac{\partial \hat{y}}{\partial b_1}\right|_o \Delta b_1 + \left.\frac{\partial \hat{y}}{\partial b_2}\right|_o \Delta b_2 + \cdots + \left.\frac{\partial \hat{y}}{\partial b_r}\right|_o \Delta b_r + D \tag{4.7}$$

or

$$\Delta \hat{y} = \left.\frac{\partial \hat{y}}{\partial b_1}\right|_o \Delta b_1 + \left.\frac{\partial \hat{y}}{\partial b_2}\right|_o \Delta b_2 + \cdots + \left.\frac{\partial \hat{y}}{\partial b_r}\right|_o \Delta b + D \tag{4.8}$$

where $\Delta\hat{y} = y\text{-}y_o$, $\Delta b_i = b_i\text{-}b_{io}$, i = 1,2,...r. The subscript (o) indicates that all quantities are evaluated at some reasonable trial value of the constants involved in each case, while D consists of the terms of higher order in Δb_i.

Assuming that the higher order terms in Δb_i are negligible, a first order approximation is given by

$$\Delta\hat{y} \doteq B_1\dot{y}_{10} + B_2\dot{y}_{20} + \cdots + B_r\dot{y}_{r0} \quad , \tag{4.9}$$

where $B_i = \Delta b_i$, $\dot{y}_{i0} = \left.\frac{\partial y}{\partial b_i}\right|_0$ and the symbol $\doteq$ means approximately equal. Now, equation (4.9) is an ordinary regression model in which the quantities $\Delta\hat{y}$ are the dependent variables, B_i are the familiar regression coefficients and the quantities $\dot{y}_{i0}$ the independent variables. The values of the latter are obtained by evaluating each partial derivative at some trial value for each parameter; the evaluations are carried out in each case for every value of x, the original independent variable considered in the model as given in equation (4.4). Estimates of the constants in this situation then are readily obtained through the direct application of ordinary regression analysis. Estimates of the variance for the parameters in the original model are also obtained simultaneously through the estimates of the variance of the B's in equation (4.9).

The B_i in equation (4.9) are corrections which are applied to the trial values of the parameters to obtain new "improved" trial values. The process is repeated successively as many times as necessary for convergence. When the system is converging, the iterative cycles should result in a net decrease in the sum of squared differences between the observed and fitted values of y. In the next section it will be shown how the Gauss-Newton method can be used for the estimation of the parameters in the model object of this dissertation.

4.2.3 Application of the Gauss-Newton Method. For the application of the Gauss-Newton method it is necessary to obtain some reasonable trial values for the parameters to be estimated; these initial values are used for starting the series of iterations. Success in approximating closely the solution values of the parameters with the trial values selected is critical for the convergence of the system. It is often possible to use prior knowledge in the selection of the necessary "educated guesses", but more often the trial values are obtained from graphs of the information available, or by solving simpler equations inherent to the system under consideration.

In the case of the proposed model for the efficiency of proteins, the trial values of the parameters for initiating the iterative procedures were obtained as follows. A graph of the same general nature as those presented in Figure 3.1 was constructed by plotting the observed responses against protein concentrations in the diet. An approximation to the maximum attainable response, η_o say, was read directly from the graph, and then used as a value of η for obtaining approximate values ξ_o and γ_o for ξ and γ respectively from the relation

$$Q = \eta - \xi e^{-\gamma p} \ . \qquad (4.10)$$

When η takes the value η_o, equation (4.10) can be written as

$$K = \xi_o e^{-\gamma_o p} \qquad (4.11)$$

where $K = (\eta_o - Q)$. Equation (4.11) can thence be linearized by taking logarithms and expressing it as

$$K' = \xi_o' - \gamma_o p \ , \tag{4.12}$$

where $K' = \ln(\eta_o - Q)$ and $\xi_o' = \ln \xi_o$; ξ_o and γ_o are the approximate solutions sought for ξ and γ, and p is the protein concentration in the diet. Estimates for ξ_o and γ_o in equation (4.12) are immediately available through the application of simple regression analysis. These are

$$\hat{\gamma}_o = \frac{\left[\Sigma\,(K'p)\right]^*}{\left[\Sigma\,(p^2)\right]^*} \ , \tag{4.13}$$

$$\hat{\xi}_o' = \bar{K}' + \gamma_o \bar{p} \ ,$$

and

$$\hat{\xi}_o = \text{antilog}\ \hat{\xi}_o' \ .$$

The symbol * indicates that the quantities in the brackets have been corrected for their means; $\bar{K}'$ and $\bar{p}$ are the arithmetic means for K' and p respectively.

A linear regression approach can also be used for obtaining estimates α_o and β_o for initiating the iterative procedure. In this case, the observed values of protein consumed at different levels of p can be used as approximations, I_o say, for the values of I, and thence estimates $\hat{\alpha}_o$ and $\hat{\beta}_o$ may be obtained by simple regression analysis from the relation $I_o = \alpha_o + \beta_o p$. These estimates are:

$$\hat{\beta}_o = \frac{\left[\Sigma\,(I_o p)\right]^*}{\left[\Sigma\,(p^2)\right]^*} \ , \tag{4.14}$$

and

$$\hat{\alpha}_o = \bar{I}_o - \beta_o \bar{p} \ .$$

Again the asterisk indicates that the quantities in brackets have been corrected for their means, and $\bar{I}_o$ and $\bar{p}$ are the arithmetic means for I_o and p respectively.

Other procedures could be used, but the one suggested has the advantage of permitting a fair degree of consistency and sufficient accuracy in determining the parameter values necessary for initiating the iterative series for obtaining the estimates of the parameters included in the model.

As it was mentioned before, not all the parameters in the model are estimable, and at least one of them must be held constant in order to estimate the others. Since reasonably fair approximations for either the maximum response (η) or the effective average food consumed (β) during the experimental period may be obtained directly from the basic data obtained in any biological trial, either one of these two parameters could be held constant through the iteration and thus, in a sense, serve as a reference point in the parameter space for the solution of the other parameters in the model. The parameter γ, which estimates the rate of curvature, is independent of the reference point chosen, but all the other parameters are affected equally by the choice in reference point since the model to be fitted is given in the form of a ratio. It should be noted that the parameter which is held constant during the iterative procedure does not have a variance. In the series of iterations which will be discussed in the next chapter, the parameter η was held constant at the value η_o, as determined for each case.

Now, from the equation for the model given in (3.21) the partial derivatives of $z=f(\underline{\theta},p)$ with respect to the parameters to be estimated are:

$$\frac{\partial \hat{z}}{\partial \xi} = -\frac{e^{-\gamma p}}{\alpha + \beta p} = \dot{y}_{10}$$

$$\frac{\partial \hat{z}}{\partial \gamma} = \frac{p\xi e^{-\gamma p}}{\alpha + \beta p} = \dot{y}_{20}$$

$$\frac{\partial \hat{z}}{\partial \alpha} = -\frac{\eta - \xi e^{-\gamma p}}{(\alpha + \beta p)^2} = \dot{y}_{30} \qquad (4.15)$$

$$\frac{\partial \hat{z}}{\partial \beta} = -\frac{(\eta - \xi e^{-\gamma p})p}{(\alpha + \beta p)^2} = \dot{y}_{40}$$

Using the initial trial values of the parameters η_o, ξ_o, γ_o, α_o and β_o the values of z_o and $\dot{y}_{io}$ $(i=1,2,3,4)$ are computed for each of the N observation points. The values $\Delta\hat{z} = (z-z_o)$ are then regressed on the $\dot{y}_{io}$ variables according to equation (4.9) to obtain the first set of estimated corrections, B_{io}, for the initial value used for the i^{th} parameter. Applying this correction to the initial values of the parameters, a new set of improved estimates ξ_1, γ_1, α_1 and β_1 is obtained; recall that η_o is to be held constant and no correction is estimated for this parameter. Now, using η_o and the improved values ξ_1, γ_1, α_1, and β_1, new values for the dependent variables, $\dot{y}_{i1}$ say, are computed and the regression procedure repeated to obtain a new set of corrections, B_{i1}, for the estimates of the parameters. The process is repeated as many times as necessary for convergence.

On approaching convergence, the correction estimates will diminish and tend to zero, the estimates of the parameters will become stable and the sum of $(\Delta\hat{z})^2$ becomes a minimum. These properties are illustrated in Figures 4.1, 4.2 and 4.3. The graphs presented were constructed using the successive results obtained in the course of a sequence of iterative cycles required for fitting the model to the results of a protein feeding trial conducted with rats. The details of this trial will be considered in the next chapter, along with others, when discussing the practical application of the model.

To obtain estimates of the corrections for the parameters, the regression coefficients in equation (4.9), it becomes necessary to solve a set of normal equations. In the case of the model under consideration with four estimable parameters, the normal equations for N sample points are given by

$$\left[Y'Y\right]\underline{B} = \left[Y'\right]\underline{z} \tag{4.16}$$

where

$$Y = \begin{bmatrix} \dot{y}_{11} & \dot{y}_{12} & \dot{y}_{13} & \dot{y}_{14} \\ \dot{y}_{21} & \dot{y}_{22} & \dot{y}_{23} & \dot{y}_{24} \\ \vdots & & & \\ \dot{y}_{N1} & \dot{y}_{N2} & \dot{y}_{N3} & \dot{y}_{N4} \end{bmatrix}, \quad \underline{B} = \begin{bmatrix} \Delta\,\xi \\ \Delta\,\gamma \\ \Delta\,\alpha \\ \Delta\,\beta \end{bmatrix} \text{ and } \underline{z} = \begin{bmatrix} \Delta z_1 \\ \Delta z_2 \\ \vdots \\ \Delta z_N \end{bmatrix} \tag{4.17}$$

The elements $\dot{y}_{ij}$ in $\left[Y\right]$ are the partial derivatives with respect to the j^{th} parameter (j=1,2,3,4) as given in (4.14), and evaluated at the i^{th}

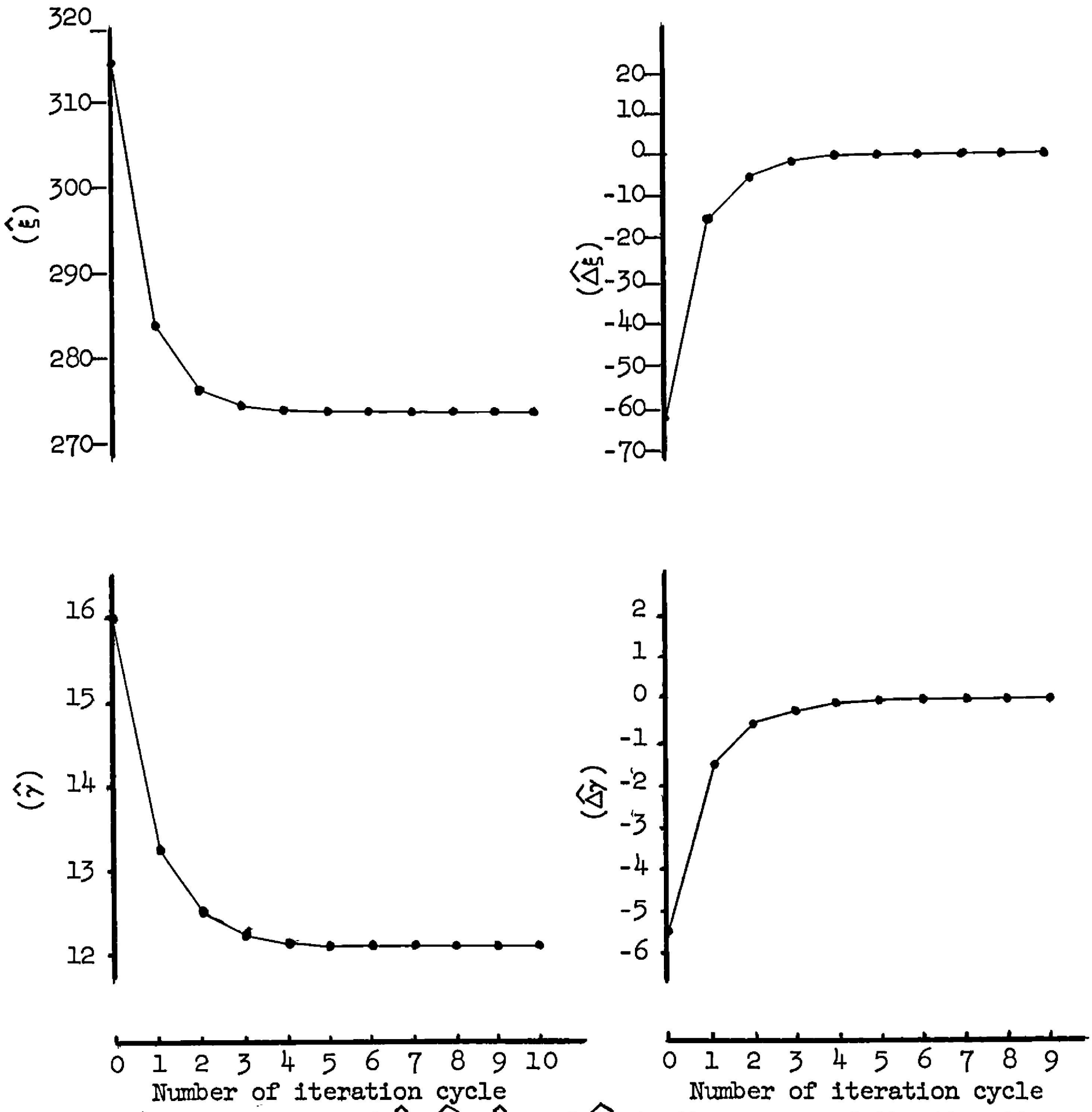


Figure 4.1. Behavior of $\hat{\xi}$, $\hat{\Delta\xi}$, $\hat{\gamma}$, and $\hat{\Delta\gamma}$ in the course of the iterative cycles necessary for the estimation of the parameters in the model for protein efficiency (Equation 3.21). The parameter η remains constant at a value η_o.

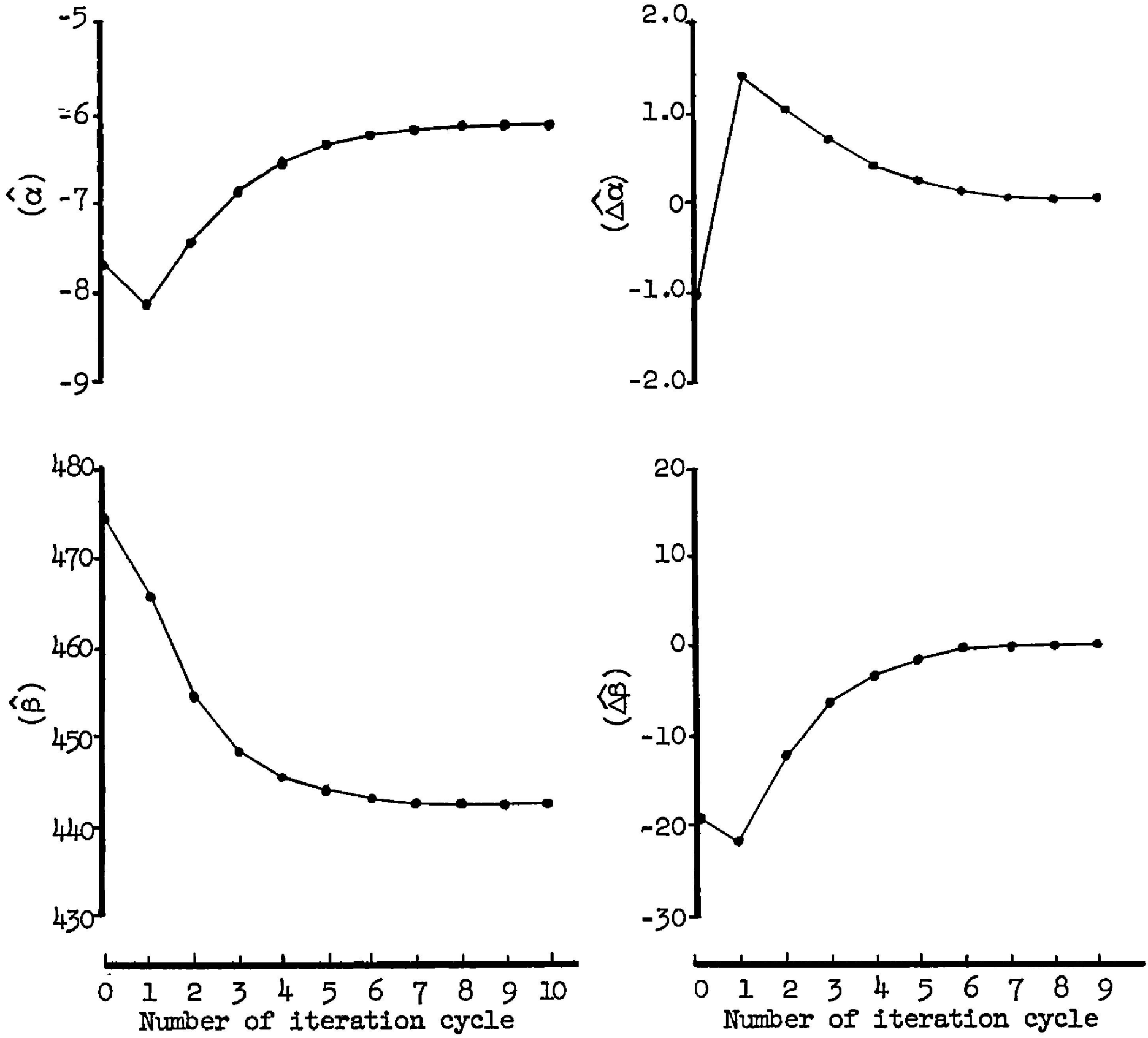


Figure 4.2. Behavior of $\hat{\alpha}$, $\hat{\Delta\alpha}$, $\hat{\beta}$, and $\hat{\Delta\beta}$ in the course of the iterative cycles necessary for estimating the parameters in the model for protein efficiency (Equation 3.21). The parameter η remains constant at a value η_o.

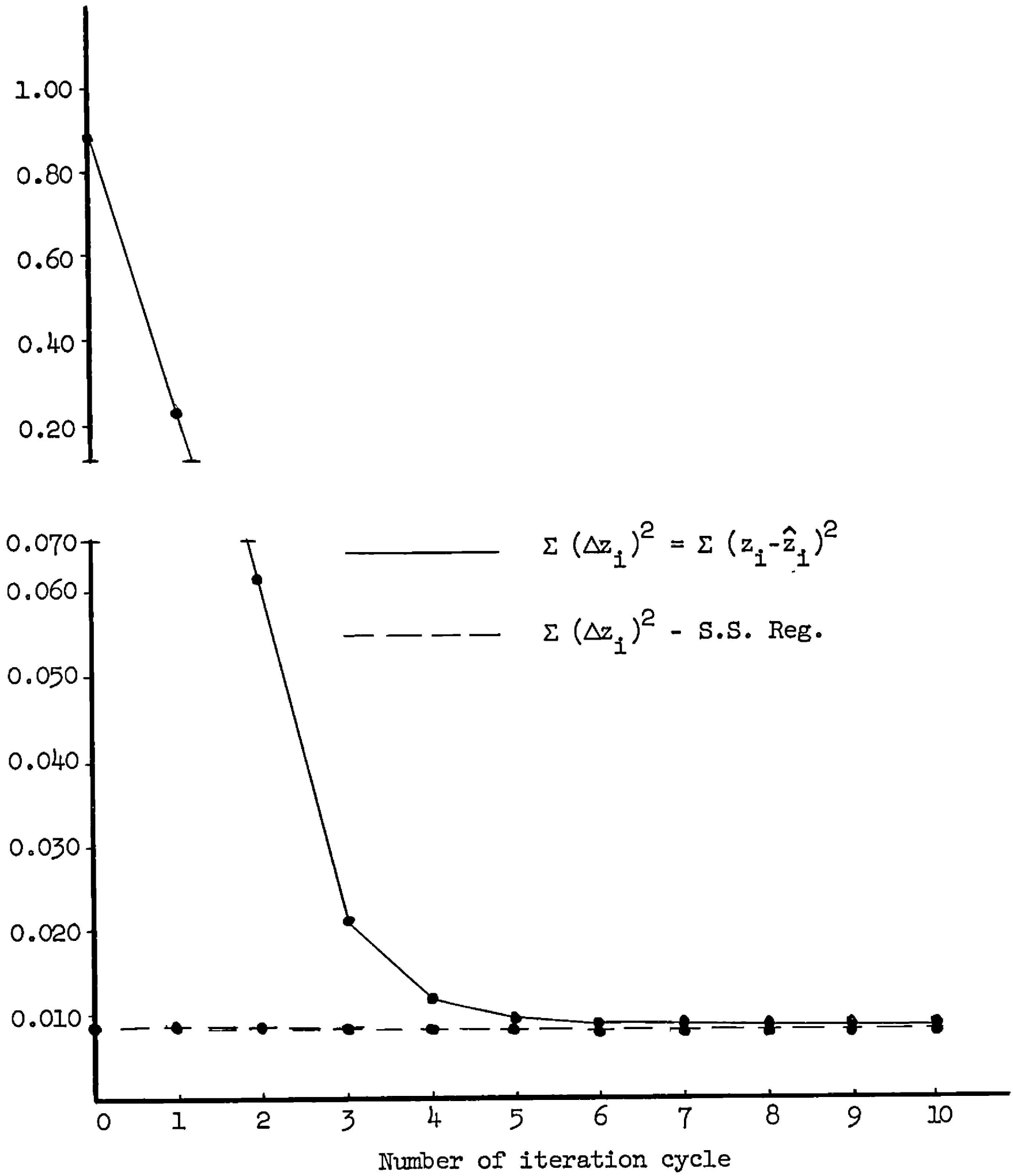


Figure 4.3. Behavior of $\Sigma (\Delta z_i)^2$ and $\left[\Sigma (\Delta z_i)^2 - \text{S.S. Regression}\right]$ in the course of the iterative cycles necessary for the estimation of the parameters in the model for protein efficiency (Equation 3.21). The parameter η remains constant at a value η_o.

sample point (i=1,2,....N) for a given set of parameter values obtained in the course of the iterative procedure. Note the change in subscript notation to conform with matrix conventions, where i is the row and j the column designation.

The solution for $\underline{B}$ in (4.16) is given by

$$\underline{B} = [Y'Y]^{-1} Y'\underline{z} \tag{4.18}$$

where $[Y'Y]^{-1}$ is the inverse of $[Y'Y]$.

The iterative procedure requires a new set of normal equations as given in (4.16) and a corresponding set of solutions as given in (4.18) for each iteration cycle. Without access to a high speed computer the task would be very tedious and time consuming. All the calculations in the present study were performed using an IBM 650, magnetic drum data-processing machine. The next section is concerned with a general comment on the program used in the iterative solution of the parameters in the model proposed in this dissertation.

4.2.4 Comment on the Program Used for the Iterations. In programming the iterative scheme for estimating the parameters in the model under consideration, an interpretive system developed at Bell Telephone Laboratories (Wolontis, 1956) was employed. This system transforms the IBM 650 into a three-address, floating decimal general purpose computer. The Bell interpretive system was chosen because it is relatively easy to use, and in addition is versatile and has a reasonably high speed of computing with full accuracy and sufficient range for elementary transcendental functions (and error less than one in the eighth digit). In fact, it is conceivable that with a minimum of

instruction and a few simple examples, a person without any previous programming experience could utilize successfully this system, and acquire sufficient experience in a short period of time to permit a fair efficiency in performance.

The program employed in this study can be outlined as follows:

1. Read in the initial values for the parameters in the model.
2. Read in the observations p_i and z_i (i=1,2,...N).
3. For the first observation point compute and store in separate locations the value of each partial derivative as given in (4.15)and the value of $z_i - \hat{z}_i$.
4. Compute and store cumulatively in separate locations the squares and crossproducts of the quantities computed in three above.
5. Change the necessary addresses of storage locations to introduce the next observation point in the series.
6. Repeat N times the operations described in three to five above, to form the matrix $[Y'Y]$ and the column vector $Y' \underline{z}$ as defined in equation (4.17).
7. Carry out the forward solution of the abbreviated Doolittle procedure (Anderson and Bancroft, 1952) and compute:
   a. The values of the regression coefficients in equation (4.16).
   b. The sum of squares due to regression and the sum of squares after regression.
8. Apply the corrections (the regression coefficients calculated in 7a) to the corresponding parameters and compute and store the resulting improved values of the parameters.

9. With the improved values of the parameters repeat the complete procedure, beginning at step three.

Optional program stops were included either to display on the IBM 650 console pertinent results at convenient intervals, and/or to introduce manually any instructions necessary for the activation of additional subroutines as desired.

In non-linear situations, curvature in the parameter space may in some cases exaggerate the estimates of the successive corrections for the parameters, often leading the iterative procedure outside of the boundaries of the region of convergence. In such cases, the introduction of certain restrictions in the iterative procedure may be helpful in preventing overly large oscillations. A procedure suggested by Wegstein (1958) and described by Turner (1959) was useful in the iterative solution for the parameters in the model under consideration. In this case, the restriction was imposed directly on the correction values by applying in each case some arbitrary fraction of the estimated corrections to the parameter values to obtain the improved values. This procedure was included in the program as a subroutine and successively decreasing or increasing fractions of the corrections were utilized as needed.

4.3 Asymptotic Variances and Confidence Limits of the Estimates

One of the advantages in employing the iterative scheme described in the previous section is related to the estimates of variance for the estimates of the parameters. The regression approach followed gives the solution for successive corrections to the parameters. In the

limit, then, the variance of the final correction estimates should be an estimate of the variance for the estimates of the parameters. In a similar situation, Turner (1959), has shown that under certain conditions the asymptotic results will suffice for the case of small samples. It should not, however, be ignored that the possibility of bias always exists for a maximum likelihood estimator.

The solution to the normal equations for the regression case under consideration was given in equation (4.18) as

$$\underline{B} = \left[\underline{Y}'\underline{Y}\right]^{-1} \underline{Y}' \underline{z} \quad .$$

In this case an estimate of σ^2 is given by

$$S^2 = \frac{1}{N-4} \left[\underline{z}'\underline{z} - \underline{B}'\underline{Y}' \underline{z}\right] \quad , \tag{4.19}$$

and then, the estimated variances for the regression coefficients, equivalent in this case to the variances of the estimates of the parameters, are given by

$$\begin{aligned} S^2_{\hat{\xi}} &= c_{11}S^2 \ , \\ S^2_{\hat{\gamma}} &= c_{22}S^2 \ , \\ S^2_{\hat{\alpha}} &= c_{33}S^2 \ , \\ S^2_{\hat{\beta}} &= c_{44}S^2 \ , \end{aligned} \tag{4.20}$$

where c_{ii} are the diagonal elements of $\left[\underline{Y}'\underline{Y}\right]^{-1}$.

Similarly, from regression theory, the variances for the predicted values of z given p can be obtained from

$$S^2_{(\hat{z}/p)} = S^2 \sum_{i=0}^{4} \sum_{i=0}^{4} c_{ij}\dot{y}_i\dot{y}_j \qquad (4.21)$$

where $\dot{y}_i$ and $\dot{y}_j$ are evaluated for the given p.

Since the estimates of the parameters are least squares estimates and equivalent to maximum likelihood estimates, they should be asymptotically normally distributed and should approach asymptotically the true values of the parameters with variances as given in equations (4.20) and (4.21). Asymptotically then, the confidence limits for the various estimates are given by

$$\hat{\xi} - t_\pi s_{\hat{\xi}} \leq \xi \leq \hat{\xi} + t_\pi s_{\hat{\xi}}$$

$$\hat{\alpha} - t_\pi s_{\hat{\alpha}} \leq \alpha \leq \hat{\alpha} + t_\pi s_{\hat{\alpha}}$$

$$\vdots \qquad (4.22)$$

$$\hat{z}/p - t_\pi s_{(\hat{z}/p)} \leq (z/p) \leq \hat{z}/p + t_\pi s_{(\hat{z}/p)}$$

with $(1-\pi)$ confidence, where t_π is student's t for the π level of significance with N-4 degrees of freedom.

CHAPTER V

APPLICATION OF THE MODEL AND DISCUSSION

5.1 Introduction

Although the assessment of protein quality has received considerable attention in the past, most of the data available have been obtained using single dietary protein concentrations, and are not adequate for testing the mathematical model proposed in this dissertation. In a few instances, however, it has been possible to use some of the published results to calculate from the observations reported the necessary basic information for fitting the model under study for a preliminary exploration of its usefulness in practice.

5.2 The Data

It was possible to express in a suitable form for fitting the model the data obtained in eleven different experiments in which seven different kinds of protein were evaluated. The experimental procedures followed, however, were not the same for all the experiments, and accordingly, some of the pertinent circumstances of each set of experiments considered will be briefly described in this section.

5.2.1 Experiments Conducted by Barnes and Associates. Barnes, et al., (1945) conducted three experiments for the evaluation of the nutritive value of the protein in whole egg (petroleum ether extracted, spray-dried), heated soyflour, and wheat gluten. In these experiments, male albino rats (Sprague-Dawley strain) of weaning age were placed in individual wire bottom cages and fed a commercial diet (Purina Fox Chow)

for three days. At the end of the pretest period, the rats were distributed according to weight in groups of eight, returned to their individual cages, and then each group was randomly assigned to one of seven different protein levels in each test diet. The test diets differed only in the source of protein they contained (for further details on the composition of the basal ration, see Barnes, et al., 1945). The protein sources replaced carbohydrates and fat in the basal ration, so that all the test diets remained isocaloric. The experimental period had a duration of 42 days, and the trials were conducted under the usual ad-libitum feeding conditions. Weight gains and food consumptions were measured throughout the experimental period, and from these observations the protein efficiency ratios (Osborne, et al., 1919) were calculated for each protein source. The data derived from the graphs of the results of these experiments as given in the original articles, are presented in Table 5.1.

5.2.2 Experiments Conducted by Bosshardt and Associates. While searching for a simplified biological procedure for the evaluation of protein quality, Bosshardt, et al., (1946) studied the protein values of casein, whole egg (coagulated, acetone extracted), and wheat gluten. These experiments were conducted with weanling albino mice (Sharp and Dohme, Swiss-Webster strain), 15-16 days of age and weighing from 7.0 to 9.0 grams. The mice were distributed according to weight in groups of seven, placed in individual wire bottom cages, and then each group randomly assigned to one of the different protein levels studied with each protein source. The test diets differed only in their source of

Table 5.1. Average weight gains and protein efficiency ratios (PER)[a] observed when feeding weanling rats for 42 days with different protein sources. Eight rats per group.

SOURCE OF PROTEIN								
Dried whole egg			Heated soyflour			Wheat gluten		
% Protein in the diet	Weight gain (grams)	PER	% Protein in the diet	Weight gain (grams)	PER	% Protein in the diet	Weight gain (grams)	PER
4.10	10	1.30	6.00	18	1.20	8.40	0	0
5.80	58	3.25	7.70	37	1.60	10.20	7	0.30
7.90	98	3.60	9.40	65	2.00	12.40	9	0.40
9.90	150	3.80	11.50	75	2.10	14.60	18	0.50
13.70	170	3.40	14.80	112	2.00	20.80	42	0.80
19.00	170	2.25	18.50	140	1.80	31.50	85	0.80
38.70	177	1.00	27.70	150	1.40	43.00	150	0.80

[a]Gain per unit of protein consumed. This table was constructed from the summary charts of the original observations as presented by Barnes, et al., (1945).

protein and the basal ration used had been shown previously to support good growth for mice (for further details on the composition of the test diets, see Bosshardt, et al., 1946). The customary procedures for ad-libitum feeding were followed during an experimental period of 20 days.

For the evaluation of casein, the protein efficiency ratios of Osborne, et al., (1919) were determined, while carcass nitrogen gain was used for the calculation of the protein efficiency of whole egg. Both weight

gains and carcass nitrogen gains were used in the case of wheat gluten. The carcass nitrogen gain was established by sacrificing a group of mice at the beginning of the experiment, determining their carcass nitrogen content and subtracting this value from the carcass nitrogen determined in the mice sacrificed at the end of the experiment in each diet group.

The protein efficiencies derived from carcass nitrogen gains were expressed as nitrogen gains per unit of nitrogen absorbed. In order to determine the absorbed nitrogen, the feces were collected throughout the experimental period and pooled for each diet group for the nitrogen determinations. The nitrogen analysis of the diets, feces and carcasses were carried out by the Kjeldahl procedure.

The results from these experiments, as derived from the charts presented by the authors, are shown in Tables 5.2 and 5.3.

5.2.3 Experiments Conducted by the Institute of Nutrition of Central America and Panama (INCAP). The Institute of Nutrition of Central America and Panama in Guatemala City, has recently developed a formula for a vegetable mixture, INCAPARINA, which through the complementation of the proteins in its ingredients results in a food with a net protein of good quality (Bressani, et al., 1961). This vegetable mixture has been intensively tested in both experimental animals and humans with satisfactory results. Data from one of these trials in which the new formula was compared with casein and dried skimmilk will be used for fitting the model.

Table 5.2. Average weight gains and protein efficiency ratios (PER)[a] observed when feeding weanling mice for 20 days with different protein sources. Seven mice per group.

SOURCE OF PROTEIN					
Casein			Wheat gluten		
% Protein in the diet	Weight gain grams	PER	% Protein in the diet	Weight gain grams	PER
5.00	1.62	0.95	--	--	--
6.25	4.10	1.40	8.12	1.11	0.35
8.50	9.00	1.78	11.69	2.20	0.50
11.22	12.60	1.70	12.50	3.49	0.59
14.69	12.30	1.30	15.50	4.41	0.61
16.94	13.90	1.15	18.31	5.56	0.65
22.89	13.60	1.05	25.50	9.41	0.70
37.50	13.50	0.70	33.56	10.41	0.72

[a]Gain per unit of protein consumed. This table was constructed from the summary charts of the original observations as presented by Bosshardt, et al., (1946).

These experiments were conducted with weanling male albino rats (Wistar strain, INCAP colony) weighing approximately fifty grams at the beginning of the trials. The rats were distributed by weight into groups of six and placed in individual wire bottom cages. Each group was then randomly assigned to one of the different levels of dietary protein furnished by either INCAPARINA, casein or dried skimmilk. The protein sources were carried isocalorically in a standard basal ration

Table 5.3. Average carcass nitrogen gains and protein efficiency ratios (PER)[a] observed when feeding weanling mice for 20 days with different protein sources. Seven mice per group.

SOURCE OF PROTEIN					
Whole egg			Wheat gluten		
% Protein in the diet	Carcass nitrogen gain, grams	PER	% Protein in the diet	Carcass nitrogen gain, grams	PER
4.12	0.774	0.43	5.75	0.072	0.05
5.44	1.302	0.51	8.12	0.258	0.09
7.44	2.034	0.52	11.69	0.444	0.10
10.81	2.376	0.38	12.50	0.660	0.13
12.62	2.382	0.32	15.50	0.864	0.14
15.25	2.472	0.29	18.31	1.080	0.14
19.62	2.712	0.26	25.50	1.728	0.15
36.62	2.412	0.12	33.56	2.016	0.15

[a]Gain per unit of nitrogen absorbed. This table was constructed from the summary charts of the original observations as presented by Bosshardt, et al., (1946).

which has proved adequate for good growth in rats; the experiments were carried out under the usual ad-libitum feeding conditions. Complete details on these trials will be published elsewhere (Bressani and Elias, 1961).

In this case, the actual observations were available in a form suitable for fitting the model. This information is given in Table 5.4.

Table 5.4. Average weight gains and protein efficiency ratios (PER)[a] observed when feeding weanling rats for 28 days with different protein sources. Six rats per group (Bressani and Elias, 1961).

SOURCE OF PROTEIN								
INCAPARINA			Casein			Dried skimmilk		
% Protein in the diet	Weight gain (grams)	PER	% Protein in the diet	Weight gain (grams)	PER	% Protein in the diet	Weight gain (grams)	PER
5.16	11	0.75	5.73	30	1.83	5.70	36	2.34
5.82	22	1.23	6.75	52	2.07	7.11	63	2.72
10.28	84	2.03	8.20	67	2.27	--	--	--
10.58	89	2.04	10.25	96	2.29	10.26	117	2.88
14.78	133	2.00	15.31	142	2.14	12.23	123	2.58
15.48	125	1.85	--	--	--	14.31	127	2.24
18.48	149	1.71	--	--	--	--	--	--
19.60	129	1.67	20.10	154	1.83	19.62	148	1.94
22.61	156	1.51	--	--	--	--	--	--
24.50	147	1.42	24.91	184	1.73	23.52	133	1.48

[a]Gain per unit of protein consumed.

5.2.4 Experiments Conducted by Bressani and Mertz. Bressani and Mertz (1958) have studied the relationship of the dietary protein level to the minimum lysine requirement of the rat. These authors employed in their studies corn gluten protein enriched with essential amino acids to match in all cases, excepting lysine, the essential amino acid content of egg protein. The enriched corn gluten protein was fed to groups

of six weanling male rats (Wistar strain, Department of Biochemistry Colony, Purdue University) for 28 days under ad-libitum feeding conditions. The effects of graded levels of lysine additions were studied at different dietary protein concentrations. (For details on the basal ration employed see Bressani and Mertz, 1958.) From the data included in the report of these authors it was possible to obtain sufficient sample points for PER for varying dietary protein concentrations at one level of lysine addition (0.80%). Unfortunately the other levels of lysine supplementation did not include sufficient dietary protein concentrations to fit the model and to permit a comparative appraisal of the results. The results from the experiment used in this dissertation are given in Table 5.5.

Table 5.5. Average weight gains and protein efficiency ratios (PER)[a] observed when feeding weanling rats for 28 days with an essential amino acid enriched corn gluten protein. Six rats per group (Bressani and Mertz, 1958).

% Protein in the diet	Weight gain grams	PER
8.00	59	3.04
12.00	119	3.15
16.00	126	2.56
20.00	125	2.00
24.00	122	1.71
32.00	115	1.26
40.00	94	0.87

[a]Gain per unit of protein consumed.

5.3 Results and Discussion

The graphs of the fitted curves for each set of data considered are presented in Figures 5.1, 5.2, 5.3, 5.4 and 5.5. The estimates of the various parameters included in the model and their standard errors for each set of data fitted are presented in Table 5.6. The medians and the geometric means of the coefficients of variability averaged over all estimates of each parameter are also included in this table. Additionally, using the estimates of the parameters, the values of some ancillary constants which have or may have some biological importance were calculated. These results are shown in Table 5.7.

5.3.1 Graphs of the Fitted Curves. Examination of Figures 5.1, 5.2, 5.3, 5.4 and 5.5 shows that in general all attempts to fit the model to the data available resulted in fairly good fits. Better fits were obtained in some cases than in others, but for all cases the behavior of the iterative procedure was as illustrated in Figures 4.1, 4.2 and 4.3 in Chapter IV using the INCAPARINA results.

The poorest fit to the model occurred with the dried whole egg data of Barnes, et al. (1945) presented in Figure 5.1. In this case, the relatively poor fit can be attributed to the odd positioning of the observation points in the region of maximum efficiency and the inability of the model to conform with the resulting distortion in the general shape of the efficiency curve. In the course of these experiments, Barnes, et al., (1945) discovered symptoms of a biotin deficiency in the rats fed the dried whole egg diets; accordingly, biotin was added to these diets at the end of the fourth week of trial. As a

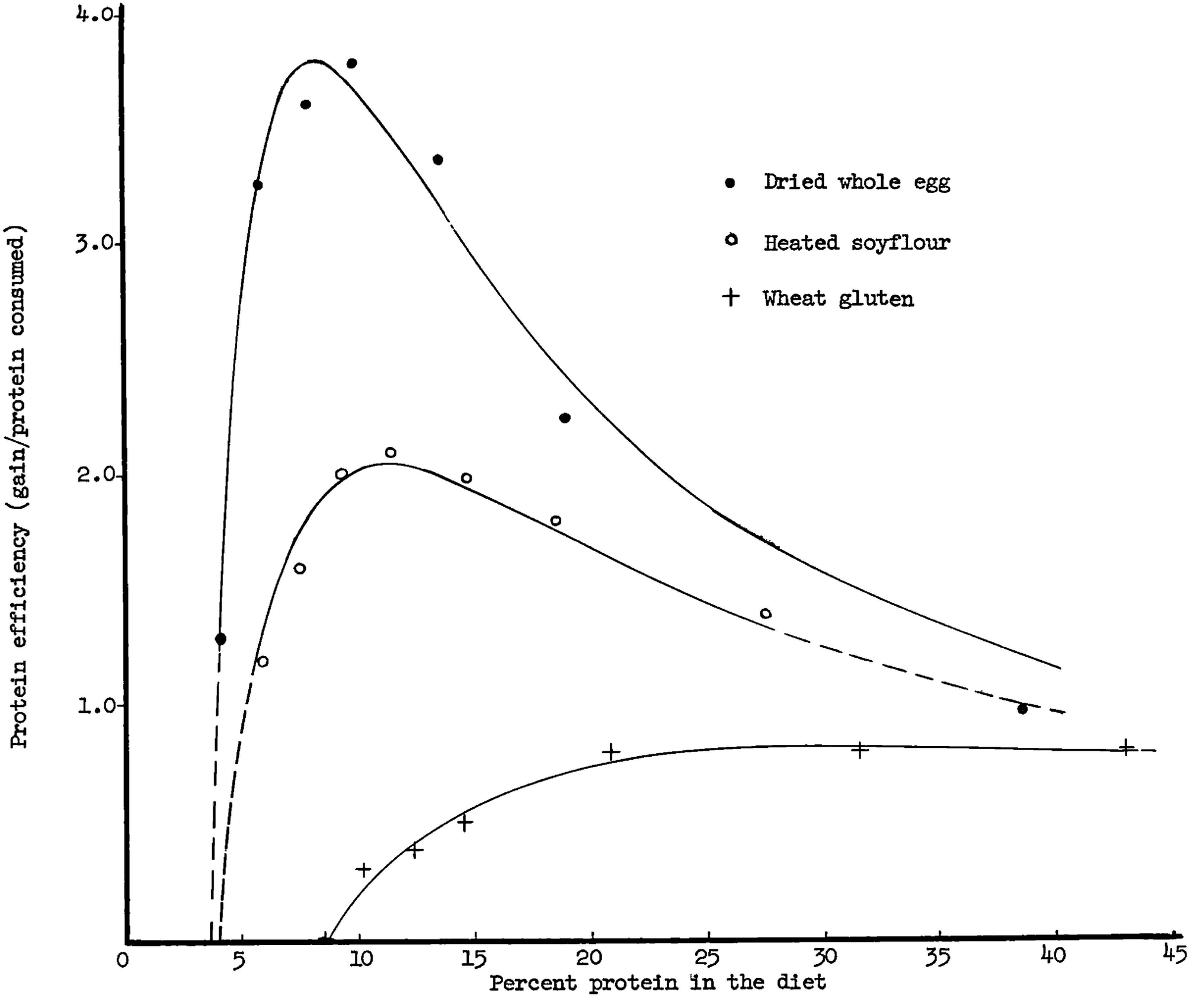


Figure 5.1. The model for protein efficiency (Equation 3.21) fitted to the data of Barnes, et al. (1945). Weight gains used as response for calculating the protein efficiencies.

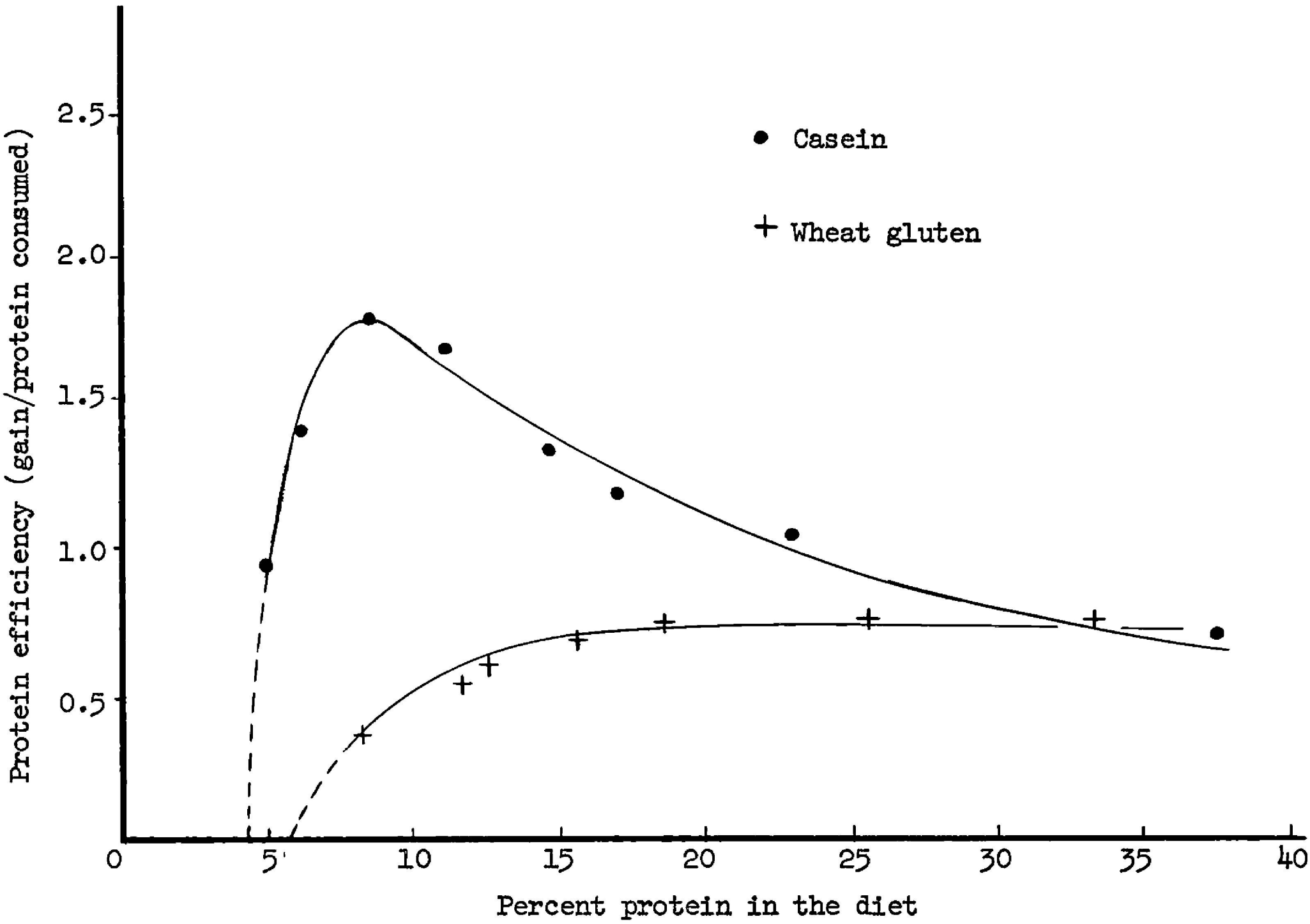


Figure 5.2. The model for protein efficiency (Equation 3.21) fitted to the data of Bosshardt, et al. (1946). Weight gains used as response for calculating the protein efficiencies.

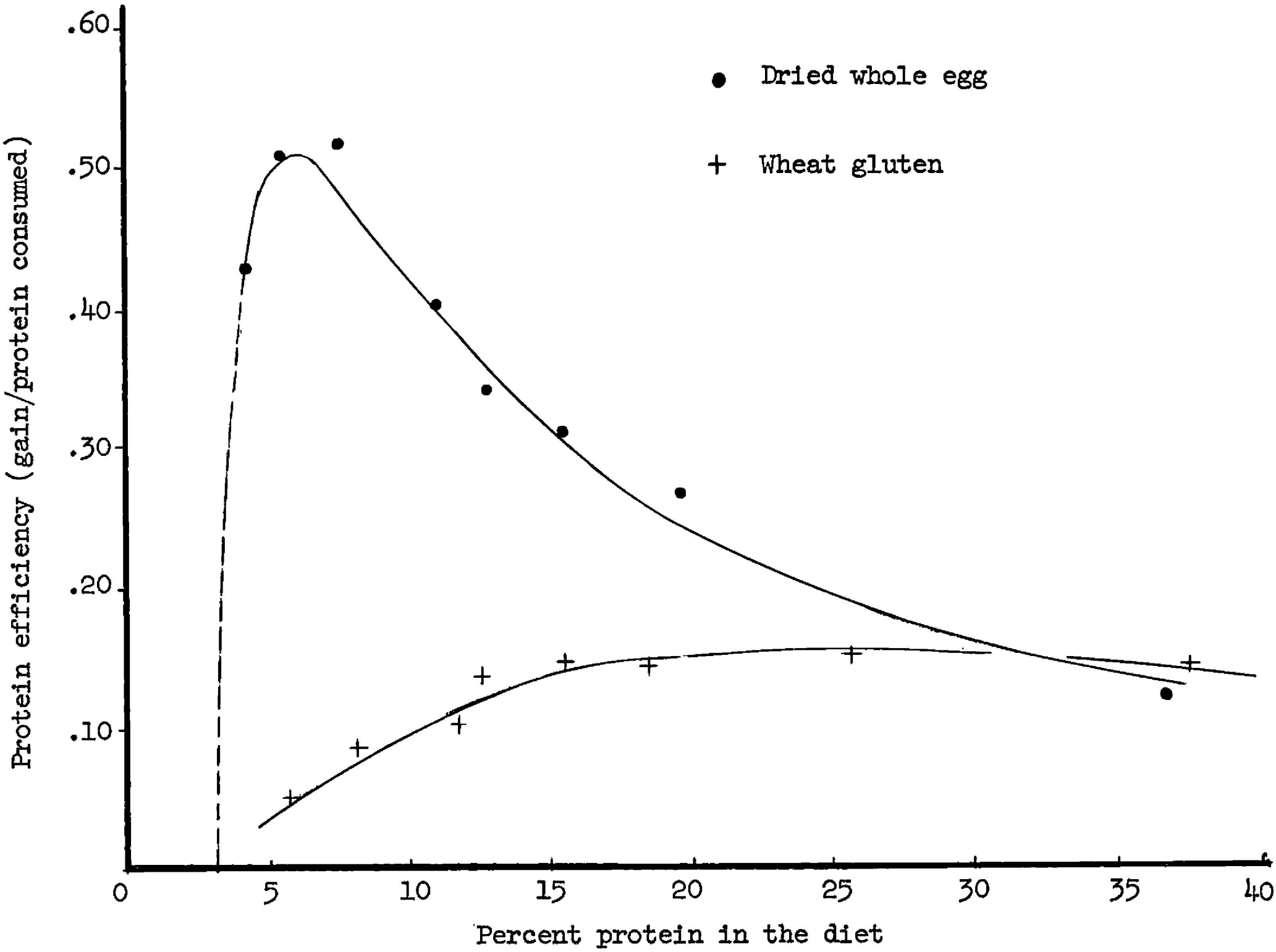


Figure 5.3. The model for protein efficiency (Equation 3.21) fitted to the data of Bosshardt, et al. (1946). Carcass nitrogen gains used as response for calculating the protein efficiencies.

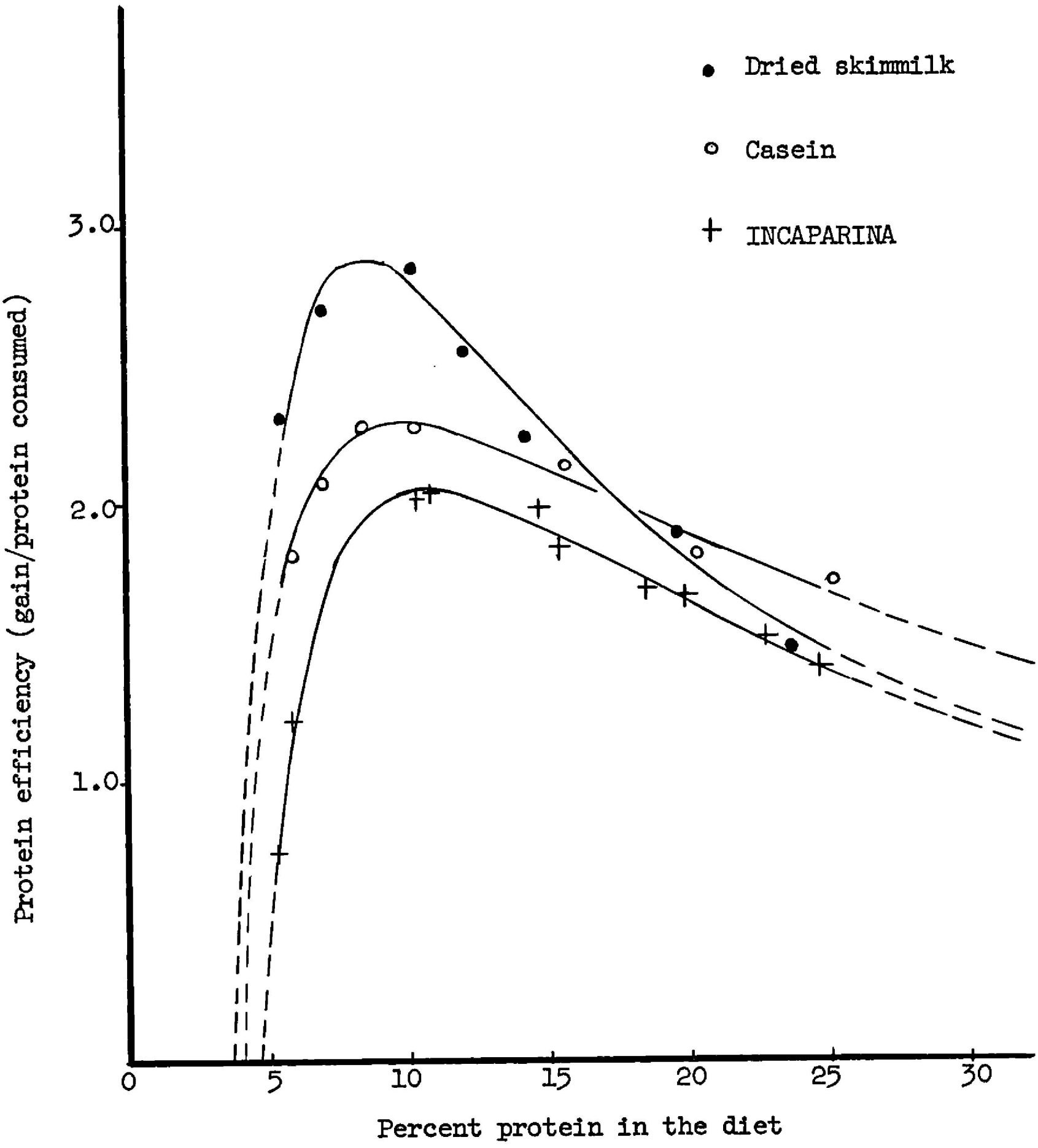


Figure 5.4. The model for protein efficiency (Equation 3.21) fitted to the INCAP data (Bressani and Elias, 1961). Weight gains used as response for calculating the protein efficiencies.

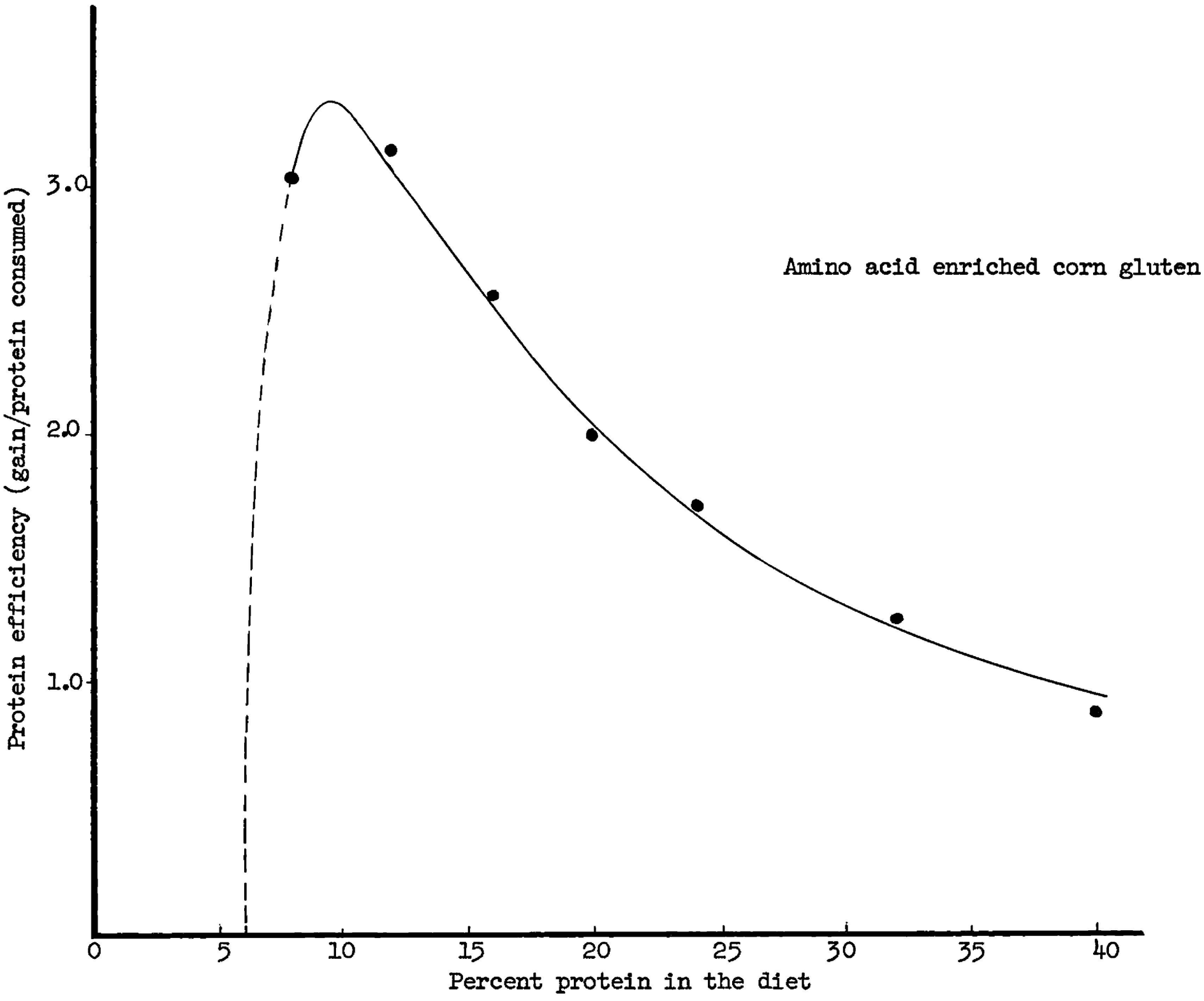


Figure 5.5. The model for protein efficiency (Equation 3.21) fitted to the data of Bressani and Mertz (1958). Weight gains used as response for calculating the protein efficiencies.

result, the rats fed the diets with lower protein concentrations (4.1 and 4.8% respectively) continued gaining weight as before, but the rats fed higher protein concentrations gained weight faster after the addition of biotin. In both cases the protein consumption remained practically unchanged, resulting in a distortion of the protein efficiencies observed. This suggests that departures from good fits may well be the result of second or higher order effects (interactions) not included in the model, and in such case the relative adequacy of the test diets in essential nutritive factors other than protein should be re-evaluated.

The shapes of the fitted efficiency curves conform with the concepts outlined in Chapter III. The proteins in foods of animal origin (whole egg, casein, and skimmilk) in general result in higher maximum efficiencies at lower dietary protein concentrations than the proteins in foods of vegetable origin (soyflour, INCAPARINA, and wheat gluten). It should be noted, however, that quality gradients occur in the two general kinds of proteins mentioned, and these are reflected in differences in the shape of the efficiency curves. Another important feature of the fitted curves pertains to the similarity between the efficiency curve obtained for corn gluten enriched with essential amino acids and the efficiency curves obtained for whole egg and skimmilk. Without amino acid enrichment, corn gluten would be expected to result in an efficiency curve similar to that of wheat gluten. The results obtained with the enriched corn gluten diets (Figure 5.5) illustrate clearly that it is possible to improve the nutritive value of proteins from vegetable sources through proper essential amino acid supplementation.

The proposed model should be particularly useful in such experiments making possible a more accurate determination of the proper level of added amino acids.

5.3.2 Estimates of the Parameters and Their Standard Errors.

With the exception of the parameter estimates for casein, which will be discussed separately, all the parameter values estimated conform with what would be expected on the basis of present knowledge concerning the quality of the proteins studied. In the discussion which follows, the estimates of all the parameters with the exception of $\hat{\gamma}$ will be considered within the framework of the particular experiment for which they were estimated; this is necessary because the estimates of maximum response (η_o), effective food consumption ($\hat{\beta}$), and hence $\hat{\xi}$ and $\hat{\alpha}$, are conditioned by the duration of the trials and the animal species used. Whenever possible, however, some effort will be made to arrive at reasonable generalizations.

The proteins of animal origin produced larger maximum responses (η_o) than the proteins of vegetable origin; the only exception occurred in the case of skimmilk which gave a slightly lower value than the one estimated for INCAPARINA. This, however, is not a surprising result for INCAPARINA is a mixture of vegetable proteins which complement their essential amino acid contents to result in a net protein of high quality. Additionally, it is not unreasonable to expect that skimmilk, when fed in high concentrations to rats, would result in conditions which would tend to limit the total food intake. It is well known, for example, that high lactose diets fed to rats produce a condition similar to diarrhea, which may limit the appetite of the animals.

Table 5.6. Parameter estimates and their standard errors.

Source of protein	η_o	Estimated values $\hat{\xi}$	$S_{\hat{\xi}}$	$\hat{\gamma}$	$S_{\hat{\gamma}}$	$\hat{\alpha}$	$S_{\hat{\alpha}}$	$\hat{\beta}$	$S_{\hat{\beta}}$
1. Experiments of Barnes and associates with rats. Weight gain used as response.									
Whole egg	177	404.24	107.94	22.96	7.18	- 0.20	6.89	374.86	22.31
Soyflour	150	340.50	30.38	19.00	1.68	14.58	6.19	347.36	11.57
Wheat gluten	150	216.43	9.81	4.23	0.59	13.24	5.74	312.92	15.03
2. Experiments of Bosshardt and associates with mice. Weight gain used as response.									
Casein	13.6	68.27	13.52	39.78	3.84	2.38	0.47	49.18	2.42
Wheat gluten	10.4	11.73	6.80	2.09	0.81	- 0.37	0.49	20.21	2.59
3. Experiments of Bosshardt and associates with mice. Carcass nitrogen gain used as response.									
Whole egg	2.7	4.97	0.23	19.02	1.33	- 1.47	0.46	61.44	4.28
Wheat gluten	2.0	2.14	0.17	3.70	0.50	3.15	1.46	46.29	6.93
4. Experiments of INCAP with rats. Weight gain used as response.									
INCAPARINA	156	268.09	20.87	12.09	1.89	- 5.96	1.52	431.41	34.35
Casein	184	244.38	27.55	7.28	1.35	- 9.15	2.21	379.86	31.28
Skimmilk	148	383.80	35.32	25.87	5.29	3.74	2.08	386.67	27.33
5. Experiments of Bressani and Mertz with rats. Weight gain used as response.									
Corn gluten + 0.80% Lys	126	402.84	52.39	19.08	2.84	-17.18	2.98	373.80	26.80
C.V.[a]-- geometric mean		11.5		13.4		49.2		7.4	
C.V.[a]--median		9.2		14.9		42.4		7.1	

[a]Coefficient of variability.

The estimates of ξ are dependent on the values of $\hat{\eta}$ and $\hat{\gamma}$ and seem to have little biological meaning other than serving as a scaling factor in the system. Note that the estimates of ξ follow a pattern similar to that for the estimates of η and γ.

The estimated values of the parameter γ appear to be relatively independent of both species and the duration of the trials. For the different proteins studied they agree with present knowledge on the quality of these proteins. The estimates of γ were generally higher for the proteins of animal origin, intermediate for heated soyflour and INCAPARINA, and low in all cases for wheat gluten. Of particular interest is the relatively high value of the estimate of γ for the enriched corn gluten protein. Without the amino acid enrichment, the value of $\hat{\gamma}$ for this protein should be similar to the value estimated for wheat gluten. This suggests that γ may well be a reasonably good index of the essential amino acid balance in the protein, and as such can be valuable in the interpretation of amino acid supplementation studies.

The estimates of γ apparently are not affected by conditions which may limit food intake; note that the value of $\hat{\gamma}$ for the skimmilk diets falls within the order of magnitude expected for this protein source and is similar to the estimate obtained for the whole egg diets.

The estimated values of α, which determine the efficiency factors for the effective food consumption, are variable and a direct interpretation of their biological meaning is difficult. It should be noted, however, that $\hat{\alpha}$ tends to be generally negative for proteins of high

quality and positive for proteins of low quality. Two exceptions, other than the erratic behavior of casein, are apparent. These occurred in the case of soyflour and skimmilk, and both of these protein sources are known to possess attributes which may limit food consumption.

The estimates of effective food consumption generally rank as expected. They are generally lower for the proteins of lower quality, probably because of a lower total calorie requirement. Soyflour and skimmilk result in lower estimates of effective food consumption than might be expected for these protein sources.

The standard errors of the estimates of the parameters presented in Table 5.6 exhibit considerable variation, and as suggested in Chapter IV may not be free of bias. However, with the exception of the estimates associated with α, the standard errors appear to be of reasonable magnitude as indicated by the geometric means and the medians of the coefficients of variability calculated over all the estimates.

5.3.3 Ancillary Constants. When functions of the parameters are used for calculating the ancillary constants presented in Table 5.7, the inconsistencies in the estimates of the parameters previously discussed seem to be minimized, and the calculated values for the constants considered agree with what would be expected on the basis of present knowledge about the quality of the protein sources studied. Proteins of better quality should result in greater efficiencies at lower dietary concentrations. Similarly, the better quality proteins should result in a balance (neither gain nor loss) with respect to response at lower

5.7. Values of ancillary constants derived from the fit to the model[a].

Source of data	Animal species	Source of protein	Response measured	Z_m	p_o	p_b
Barnes, et al., 1945	rat	dried whole egg	weight gain	3.78	0.0823	0.0359
Bosshardt, et al., 1946	mice	dried whole egg	carcass N gain	0.50	0.0590	0.0318
Bosshardt, et al., 1946	mice	casein	weight gain	2.87	0.0863	0.0405
Bressani and Elias, 1961	rat	casein	weight gain	2.28	0.0950	0.0392
Bressani and Elias, 1961	rat	skimmilk	weight gain	2.88	0.0850	0.0368
Barnes, et al., 1945	rat	heated soy flour	weight gain	2.04	0.1168	0.0379
Bressani and Elias, 1961	rat	INCAPARINA	weight gain	2.05	0.1075	0.0448
Barnes, et al., 1945	rat	wheat gluten	weight gain	0.84	0.2924	0.0878
Bosshardt, et al., 1946	mice	wheat gluten	weight gain	0.74	0.2350	0.0512
Bosshardt, et al., 1946	mice	wheat gluten	carcass N gain	0.16	0.2045	0.0157
Bressani and Mertz, 1958	rat	enriched corn gluten	weight gain	3.30	0.0961	0.0608

[a] Z_m = maximum efficiency; p_o = optimum protein concentration, i.e., that concentration at which the maximum efficiency occurs; p_b = protein concentration for "balance" or "maintenance", i.e., that concentration of protein at which neither gains nor losses are observed.

dietary protein concentrations than the proteins of lesser quality. With proteins of very poor quality the results may be variable because, in these cases, neither the maximum efficiency nor the corresponding dietary protein concentration are sharply defined. Similarly, the responses observed at low dietary concentrations of poor quality proteins can be expected to be very variable, and therefore, the value of the protein concentration required for balance may also be poorly determined. The results presented in Table 5.7 for the different proteins studied agree well with the concepts outlined above. Additionally, these results seem to suggest that the values of these constants are fairly independent of species; the results for the maximum efficiency index are, of course, dependent on the response criteria employed, but can be easily converted to a common scale. In the present case, the results obtained when using carcass nitrogen gain as a response are equivalent to those obtained when weight gain was used as response, making the proper correction for water and fat content of the body.

5.3.4 Remarks on the Casein Experiments. The erratic behavior of the estimates of the parameters for casein (Table 5.6) is not biologically impossible. Casein is known to be variable in quality and capable of producing extreme results when biologically tested. The procedures followed in the processing of this product can be critical and may affect its nutritive value. Different grades of casein covering a wide range of purity specification are available in the common market. The apparently contradictory results in the parameter estimates in the two casein experiments could be attributed to the use of two casein products of different grades of purity. There are, however,

some peculiarities in the basic data used to fit the model, which suggest that factors other than differences in purity may contribute to the observed discrepancies.

In the INCAP data on casein (Bressani and Elias, 1961) presented in Table 5.5, an added response in weight gain seems to occur when the diet contains 25% or more protein. This is not an unlikely result. Casein is known to be limiting in cystine-methionine (Mitchell, 1959), tryptophan and isoleucine (Frost, 1959) in that order of importance. If the relative excesses of the other amino acids with respect to the limiting factors are not detrimental to the animal, a response increment can be expected to occur at the higher levels of dietary protein. This response increment would tend to diminish the rate of decrease in efficiency, if the food intake remains essentially constant, damping the curvature of the system. This condition would be reflected in a low value of γ, as was the case with the estimate derived from the INCAP casein experiment (Table 5.6).

The results from the casein experiment conducted by Bosshardt, et al., (1946) cannot be explained in the same manner. In this case, either the authors used a casein of a much better grade of purity, or else random fluctuations in the observations at the critical point in response curvature resulted in an artificially high value for the estimate of γ, with a corresponding increase in the value of $\hat{\xi}$. It should be recalled that in this case the original data were not available for the fitting and the sample points used had to be calculated from the summary graphs in the published reports. Unavoidable errors are

introduced in using this approximate procedure. In this connection, the data presented in Table 5.2 suggest that either one or two observations ($p = 14.69$ and $p = 16.94$) may be out of line and may be responsible for the high value of the estimates of γ and ξ. The estimate of γ, which is determined by the curvature in the system, can be affected by errors in the observations, specially when these occur near or on the region of maximum curvature for the response considered. This effect can be particularly marked when the fitting is done on the basis of a small number of sample points. In this case it was considered advisable to use an alternate estimating procedure (Stevens, 1951) for fitting the numerator of the model alone and verify the results obtained when fitting the complete model. The results given by the alternate method of fitting were essentially the same as before and also resulted in a discrepancy in the estimates of γ and ξ of the same order of magnitude as reported in Table 5.6 for the two casein experiments.

5.4 General Comment

The proposed model for the nutritional evaluation of proteins gives useful information for a better understanding of the factors which determine protein quality. The results described in the previous section, however, suggest that additional research is necessary for the improvement and better characterization of the system.

Other methods of fitting the model need to be examined to increase the efficiency of the estimates and of the experimental errors. The biases that may be involved in the various estimation procedures should

be studied by empirical sampling methods which consider both additive and multiplicative errors. This information would be valuable for the development of optimum experimental designs for the assessment of protein quality.

The model should be extended to allow a better quantification of the effect which non-protein factors and the use of experimental periods of different duration may have in the evaluation of protein quality. A better characterization of "toxic" effects should be introduced and other parameterizations which may be more nearly species independent and/or free of the effect of interactions between nutritive entities should be developed.

Finally, by devising suitable cost functions which can be related to the estimates of the parameters and to the values of the ancillary constants that may be calculated from these estimates, it should be possible to give attention to pertinent economical factors to permit a more efficient utilization of available protein sources.

CHAPTER VI

SUMMARY AND CONCLUSIONS

Awareness of the critical importance of proper protein nutrition to a large sector of the world population, has stimulated in recent years interest in the search for new and cheaper sources of protein foods. As a natural consequence of these activities, it is becoming increasingly evident that better methods for the evaluation of protein quality are needed.

Most of the procedures commonly employed in the evaluation of protein quality are based on the assessment of the relative efficiency of different proteins in promoting observable biological responses. Additionally, some attempts have been made to predict the biological efficiencies of proteins on the basis of their amino acid make-up. Although the latter attempts have given some promising results, they cannot as yet be considered sufficiently reliable. Accordingly, the biological procedures first mentioned are usually preferred for the assessment of protein quality.

The biological methods used for the evaluation of protein quality differ primarily in the choice of response measured and in the experimental procedure followed. In spite of these differences, however, they all ultimately result in essentially the same general order of quality rankings for most food proteins. This similarity in results suggests that it should be possible to formulate a general model for the evaluation of protein quality. Additionally, if in such a model the responses observed are evaluated on a per unit of protein input basis (efficiencies),

since cost may be assumed to be proportional to food and protein consumptions, it should be possible to consider important economic factors.

This dissertation is concerned with the development and application of a mathematical model which relates the protein concentration in the diet to the observed efficiencies. In developing the model protein induced responses were assumed to obey a law of diminishing returns with respect to the concentration of protein in the diet, while the rate of change of protein input into the system was assumed to remain constant with respect to the dietary protein concentration. Accordingly, observed efficiencies (z) can be expressed as

$$z = \frac{\eta - \xi e^{-\gamma p}}{\alpha + \beta p} ,$$

where η is the average maximum attainable response for all individuals of a given species; ξ is a scaling factor which varies with η and γ; γ is an index of protein quality; α is a general correction term for the general adequacy of the diet; β is a measure of "effective" food consumption, proportional to the observed food consumption; and p is the protein concentration in the diet. Of these parameters α and γ seem to be independent of species, while η, ξ and β depend on the animal species used in the feeding trials.

This model permits the estimation of at least three important attributes of protein quality, and should allow a better understanding of the meaning of protein quality. In this connection, the joint consideration of the parameters γ and η might prove valuable for the understanding of amino acid unbalance.

Three ancillary constants which are independent of the species used in the trials can be calculated from the estimates of the parameters. A simple graphic procedure may be used for estimating the optimum dietary protein concentration (p_o), the dietary protein concentration necessary for maximum efficiency. The estimate of p_o can then be applied to the model to obtain the value of the maximum efficiency (Z_m). The protein concentration at which neither gains nor losses occur (p_b) can be calculated from the simple relation

$$p_b = \frac{1}{\gamma} \ln \left[\frac{\xi}{\eta} \right] .$$

The three constants p_o, Z_m, and p_b also reflect protein quality, but additionally, because of their nature, should facilitate the introduction of economical consideration in relation to protein quality.

For the application of the model, the proteins tested must be fed over a range of protein concentrations. This procedure eliminates the usual discrepancies in results when the biological trials are conducted using different single protein concentrations. This feature of the model made somewhat difficult its practical application because experiments covering a sufficient number of dietary protein concentrations are not abundant. However, some data obtained with either mice or rats by different investigators, using either weight gains or carcass nitrogen gains as response criteria, were available in the literature and were used for fitting the model. The protein sources covered in these trials included dried whole egg, skimmilk, casein, amino acid enriched corn gluten, heated soyflour, INCAPARINA and wheat gluten. The results from these fittings indicate that the proposed model for the nutritional

evaluation of proteins gives useful information for a better understanding of the factors which determine protein quality. The results described, however, suggest that additional research is necessary for the improvement and better characterization of the system.

Other methods of fitting the model need to be examined to increase the efficiency of the estimates and of the experimental errors. The biases that may be involved in the various estimation procedures should be studied by empirical sampling methods which consider both additive and multiplicative errors. This information would be valuable for the development of optimum experimental designs for the assessment of protein quality.

The model should be extended to allow a better quantification of the effect which non-protein factors and the use of experimental periods of different duration may have in the evaluation of protein quality. A better characterization of "toxic" effects should be introduced and other parameterizations which may be more nearly species independent and/or free of the effect of interactions between nutritive entities should be developed.

Finally, by devising suitable cost functions which can be related to the estimates of the parameters and to the values of the ancillary constants that may be calculated from these estimates, it should be possible to give attention to pertinent economical factors to permit a more efficient utilization of available protein sources.

LIST OF REFERENCES

Allison, J. B. 1953. Amino acid requirements of man. Borden's Rev. Nutrition Research 14:61-79.

Allison, J. B. 1954. Other methods of evaluation in Spector, H., Peterson, M. S. and Friedman, T. E. (Eds.) Methods for Evaluation of Nutritional Adequacy and Status. National Academy of Sciences, National Research Council, Washington.

Allison, J. B. 1955. Biological evaluation of proteins. Physiol. Revs. 35:664-700.

Allison, J. B. 1957. Nitrogen balance and the nutritive value of proteins. J. Am. Med. Assoc. 164:283-289.

Allison, J. B. and Anderson, J. A. 1945. The relation between absorbed nitrogen, nitrogen balance and biological value of proteins in adult dogs. J. Nutrition 29:413-420.

Allison, J. B., Anderson, J. A. and Seeley, R. D. 1946. The determination of the nitrogen balance index in normal and hypoproteinemic dogs. Ann. N. Y. Acad. Sci. 47:245-271.

Almquist, H. J. 1953. Interpretation of amino acid requirement data according to the law of diminishing returns. Arch. Biochem. Biophys. 44:245-247.

Almquist, H. J. and Merritt, J. B. 1952. Effect of raw soybean meal on growth of the chick. Proc. Soc. Exp. Biol. Med. 79:277-279.

Anderson, R. L. and Bancroft, T. A. 1952. Statistical Theory in Research. McGraw-Hill Book Co., Inc., New York.

Barnes, R. H. and Bosshardt, D. K. 1946. The evaluation of protein quality in the normal animal. Ann. N. Y. Acad. Sci. 47:273-296.

Barnes, R. H., Maack, Jean E., Knights, Mary J., and Burr, G. O. 1945. The measurement of the growth promoting quality of dietary protein. Cereal Chem. 22:273-286.

Béhar, M., Viteri, F., Bressani, R., Arroyave, G., Squibb, R. L., and Scrimshaw, N. S. 1958. Principles of treatment and prevention of severe protein malnutrition in children (Kwashiorkor). Ann. N. Y. Acad. Sci. 69:954-968.

Block, R. J. 1956. The protein requirements of animals including men. Borden's Rev. Nutrition Research 17:75-96.

Block, R. J. and Mitchell, H. H. 1946. The correlation of the amino acid composition of proteins with their nutritive value. Nutrition Abstr. and Revs. 16:249-278.

Bosshardt, D. K., Ydse Laura, C., Ayres, Mary M. and Barnes, R. H. 1946. The use of mice for the measurement of the growth promoting quality of proteins. J. Nutrition 31:23-33.

Boussingault, J. B. 1836. Recherches sur la quantité d'azote contenu dans les fourrages, et sur leurs équivalents. L'Académie des Sciences, Paris. Comptes Rendus 3:726-732.

Boussingault, J. B. 1838. Analyses comparées des aliments consommés et des produits rendus par un cheval, soumis a la ration d'avoines; suite des recherche enterprises dans le but d'examiner si les herbivores empruntent de l'azote à l'atmosphére. L'Académie des Sciences, Paris. Comptes Rendus 7:1157-1160.

Braham, J. E., Bressani, R. and Guzmán, M. A. 1959. A rapid procedure for the determination of net protein utilization (NPU) with New Hampshire chicks (Abstract). Federation Proc. 18:518.

Bressani, R. 1960. Unpublished data. Instituto de Nutrición de Centro América y Panamá, Guatemala, C. A.

Bressani, R. and Elias, L. G. 1961. All-vegetable protein mixtures for human feeding. VI. Biological testing of INCAP vegetable mixture nine in rats. (Manuscript in preparation.)

Bressani, R. Elias, L. G., Aguirre, A. and Scrimshaw, N. S. 1961. All-vegetable protein mixtures for human feeding. III. Development of INCAP vegetable mixture nine. J. Nutrition 74:201-208.

Bressani, R. and Mertz, E. T. 1958. Relationship of protein level to the minimum lysine requirement of the rat. J. Nutrition 65:481-492.

Bressani, R., Scrimshaw, N. S., Béhar, M. and Viteri, F. 1958. Supplementation of cereal proteins with amino acids. II. Effect of amino acid supplementation of corn-masa at intermediate levels of protein intake on the nitrogen retention of young children. J. Nutrition 66:501-513.

Bressani, R., Wilson, D., Béhar, M. and Scrimshaw, N. S. 1960. Supplementation of cereal proteins with amino acids. III. Effect of amino acid supplementation of wheat flour as measured by nitrogen retention of young children. J. Nutrition 70:176-186.

Brock, J. F. and Autret, M. 1952. Kwashiorkor in Africa. World Health Organization, Monograph Series No. 8, WHO, Geneva.

Cannon, P. R., Humphreys, Eleanor M., Wissler, R. W., and Frazier, L. E. 1944. Chemical, clinical and immunological studies on the products of human plasma fractionation. XXIII. The effects of feeding possible blood substitutes on serum protein regeneration and weight recovery in the hypoproteinemic rat. J. Clin. Invest. 23:601-606.

Chanutin, A. and Mendel, L. B. 1922. A comparison of the nitrogenous metabolism during single and fractional feeding. J. Metabolic Research 1:481-488.

Crampton, E. W. and Lloyd, L. E. 1959. Fundamentals of Nutrition. W. H. Freeman and Co., San Francisco.

Downes, Helen R. 1955. The Chemistry of Living Cells. Harper, New York.

Elvehjem, C. A., and Krehl, W. A. 1955. Dietary interrelationships and imbalance in nutrition. Borden's Rev. Nutrition Research 16:69-84.

FAO. 1957. Protein Requirements. Report of the Food and Agriculture Organization of the United Nations Committee on Protein Requirements. Nutritional Studies No. 16. FAO, Rome.

Fingerling, G. 1937. Ersatz des Nahrungseiweisses durch Harnstoff beim wachsenden Rinde. Landwirtsch. Vers.-Sta. 128:221-235.

Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. Phil. Trans. Roy. Soc., London Ser. A. 222:309-368.

Flodin, N. W. 1957. Amino acid balance and efficiency of protein utilization. Metabolism, Clin. and Exp. 6:350-364.

Frazier, L. E., Wissler, R. W., Steffee, C. H., Woolridge, R. L. and Cannon, P. R. 1947. Studies in amino acid utilization. I. The dietary utilization of mixtures of purified amino acids in protein-depleted adult albino rats. J. Nutrition 33:65-84.

Fruton, J. S. and Simmonds, Sofia. 1959. General Biochemistry. Second Edition. John Wiley and Sons, Inc., New York.

Frost, D. V. 1959. Methods of measuring the nutritive value of proteins, protein hydrolyzates, and amino acid mixtures. The repletion method. in Albanese, A. A. (Ed.) Protein and Amino Acid Nutrition. Academic Press, New York.

Folin, O. 1905. A theory of protein metabolism. Am. J. Physiol. 13: 117-138.

Gandra, Y. R. and Scrimshaw, N. S. 1961. Infection and nutritional status. II. Effect of mild virus infection induced by 17-D yellow fever vaccine on nitrogen metabolism in children. Am. J. Clin. Nutrition 9:159-163.

Gómez, F., Ramos-Galván, R., Gravioto, J., and Frenk, S. 1958. Prevention and treatment of chronic severe infantile malnutrition (Kwashiorkor). Ann. N. Y. Acad. Sci. 69:969-981.

Guyton, A. C. 1956. Textbook of Medical Physiology. W. B. Saunders and Co., Philadelphia.

Guzmán, M. A., Ascoli, W., and Scrimshaw, N. S. 1961. Nutrition research in Spanish-Portugese speaking countries in Brock, J. F. Recent Advances in Nutrition with Special Reference to Clinical Medicine. J. and A. Churchill, Ltd., London.

Harper, A. E. 1958. Balance and imbalance of amino acids. Ann. N. Y. Acad. Sci. 69:1025-1038.

Hundley, J. M. 1959. Malnutrition - a global problem. Federation Proc. 18:76-81.

INCAP. 1961. Unpublished data. Instituto de Nutrición de Centro América y Panamá, Guatemala, C. A.

Kühnau, J. 1949. Biochemie des Nahrungseiweisses. Angew. Chem. 61: 357-365.

Leverton, Ruth M., Gram. Mary R., Chaloupka, Marilyn, Brodovsky, Eileen, and Mitchell, Amy. 1956. The quantitative amino acid requirements of young women. I. Threonine. J. Nutrition 58:59-81.

Magendie, F. 1816. Sur le propriétés nutritives des substances qui ne contiennent pas d'azote. Ann. Chim. Phys. 1st. Ser. 3:66-77.

Magendie, F. 1841. Rapport fait à l'Académie des Sciences au nom de la Commission dite de la gélatine. L'Académie des Sciences, Paris. Comptes Rendus 13:237-283.

Maynard, L. A. 1951. Animal Nutrition. Third Edition. McGraw-Hill Book Co., Inc., New York.

Melnick, D. and Cowgill, G. R. 1937. The protein minima for nitrogen equilibrium with different proteins. J. Nutrition 13:401-424.

Miller, D. S. and Bender, A. E. 1955. The determination of the net utilization of proteins by a shortened method. Brit. J. Nutrition 9:382-388.

Mitchell, H. H. 1923-24a. A method of determining the biological value of protein. J. Biol. Chem. 58:873-903.

Mitchell, H. H. 1923-24b. The biological value of proteins at different levels of intake. J. Biol. Chem. 58:905-922.

Mitchell, H. H. 1923-24c. The supplementary relations among proteins. J. Biol. Chem. 58:923-929.

Mitchell, H. H. 1943. Biological methods of measuring the protein values of feeds. J. Animal Sci. 2:263-277.

Mitchell, H. H. 1944. Determination of the nutritive value of the proteins of food products. Ind. Eng. Chem., Anal. Ed. 16:696-700.

Mitchell, H. H. 1954. Biological value of proteins and amino acid interrelationships in Spector, H., Peterson, M. S. and Friedman, T. E. (eds.) Methods for Evaluation of Nutritional Adequacy and Status. National Academy of Sciences, National Research Council, Washington.

Mitchell, H. H. 1955. The validity of Folin's concept of dichotomy in protein metabolism. J. Nutrition 55:193-208.

Mitchell, H. H. 1959. Some species and age differences in amino acid requirements in Albanese, A. A. (Ed.) Protein and Amino Acid Nutrition. Academic Press, New York.

Mitchell, H. H. and Block, Richard J. 1946. Some relationships between the amino acid contents of proteins and their nutritive values for the rat. J. Biol. Chem. 163:599-620.

Mitchell, H. H. and Carman, G. G. 1924. The biological value for maintenance and growth of the proteins of whole wheat, eggs, and pork. J. Biol. Chem. 60:613-620.

Mitchell, H. H. and Hamilton, T. S. 1949. The dermal excretion under controlled environmental conditions of nitrogen and minerals in human subjects, with particular reference to calcium and iron. J. Biol. Chem. 178:345-361.

Mitchell, H. H., Hamilton, T. S. and Beadles, J. R. 1945. The importance of commercial processing for the protein value of food products. I. Soybean, coconut and sunflower seed. J. Nutrition 29:13-25.

Monroe, R. J. 1949. On the use of non-linear systems in the estimation of nutritional requirements of animals. Unpublished Ph.D. Thesis, North Carolina State College, Raleigh.

Mulder, G. J. 1839. Ueber die Zusammensetzung einiger thierischen Substanzen. J. Prakt. Chem. 16:138. Reproduced as photostat of original article in Schmidt, C. L. A. 1945. The Chemistry of the Amino Acids and Proteins. Charles C. Thomas, Springfield, Ill.

Murlin, J. R., Nasset, E. S., and Marsh, M. Elizabeth. 1938. The egg-replacement value of the proteins of cereal breakfast foods with a consideration of heat injury. J. Nutrition 16:249-269.

N.R.C. 1959. Evaluation of Protein Nutrition. A report of the Food and Nutrition Board, Division of Biology and Agriculture. Publication 711. National Academy of Sciences, National Research Council, Washington.

Osborne, T. B. 1907. The Proteins of the Wheat Kernel. Publication 84. Carnegie Institution of Washington.

Osborne, T. B. and Mendel, L. B. 1917. The use of soybean as food. J. Biol. Chem. 32:369-387.

Osborne, T. B., Mendel, L. B. and Ferry, Edna L. 1919. A method for expressing numerically the growth promoting value of proteins. J. Biol. Chem. 37:223-229.

Oser, B. L. 1951. Method for integrating essential amino acid content in the nutritional evaluation of protein. J. Am. Dietet. Assoc. 27:396-402.

Perez, C. 1959. Progress in understanding and preventing protein malnutrition in Central America. Federation Proc. 18:89-93.

Platt, B. S. and Miller, D. S. 1959. The net dietary-protein value (N.D.-p.v.) of mixture of foods - its definition, determination and application. Proc. Nutrition Soc. 18:vii-viii.

Reid, J. T. 1953. Urea as a protein replacement for ruminants: A review. J. Dairy Sci. 36:955-996.

Rose, W. C. 1937. The nutritive significance of the amino acids and certain related compounds. Science 86:298-300.

Rose, W. C. 1938. The nutritive significance of the amino acids. Physiol. Revs. 18:109-136.

Scarborough, J. B. 1950. Numerical Mathematical Analysis. Johns Hopkins Press, Baltimore.

Schneider, B. H. 1935. The subdivision of the metabolic nitrogen in the feces of the rat, swine and man. J. Biol. Chem. 109:249-278.

Schoenheimer, R., Ratner, S., and Rittenberg, D. 1939. Studies in protein metabolism. X. The metabolic activity of body proteins investigated with ℓ(-)-leucine containing two isotopes. J. Biol. Chem. 130:703-732.

Scrimshaw, N. S. 1961. Nutrition and infection. in Brock, J. F. Recent Advances in Nutrition with Special Reference to Clinical Medicine. J. and A. Churchill, Ltd., London.

Scrimshaw, N. S. and Béhar, M. 1959. World-wide occurrence of protein malnutrition. Federation Proc. 18:82-88.

Scrimshaw, N. S. and Béhar, M. 1961. Protein malnutrition in young children. Science 133:2039-2047.

Scrimshaw, N. S., Taylor, C. E. and Gordon, J. E. 1959. Interactions of nutrition and infection. Am. J. Med. Sci. 237:367-403.

Sebrell, W. H. and Hand, D. B. 1957. Protein malnutrition as a world problem, in Amino Acid Malnutrition. Rutgers University Press, New Brunswick, N. J.

Sénécal, J. 1958. The treatment and prevention of Kwashiorkor in French West Africa. Ann. N. Y. Acad. Sci. 69:916-953.

Stevens, W. L. 1951. Asymptotic regression. Biometrics 7:247-267.

Thomas, Karl. 1909. Über die biologische Wertigkeit der Stickstoffsubstanzen in verschiedenen Nahrungsmitteln. Arch. Anat. Physiol., Physiol. Abt. 219-302.

Thompson, D'Arcy Wentworth. 1942. On Growth and Form. A new edition. The University Press, Cambridge, England.

Tukey, J. W. n.d. 1958? The propagation of errors, fluctuations and tolerances: basic generalized formulas. Princeton University, Department of Mathematics, Section of Mathematical Statistics, Statistical Techniques Research Group. Technical Report No. 10., Princeton, N. J.

Tukey, J. W. n.d. 1960? The propagation of errors, fluctuation and tolerances. 3. An exercise in partial differentiation. Princeton University, Department of Mathematics, Section of Mathematical Statistics, Statistical Techniques Research Group. Technical Report No. 12., Princeton, N. J.

Turner, M. E. 1959. The single process law: a study in nonlinear regression. Unpublished Ph.D. Thesis, North Carolina State College, Raleigh.

Vickery, H. B. and Osborne, T. B. 1928. A review of hypotheses of the structure of proteins. Physiol. Revs. 8:393-446.

Vivanco, F. 1960. Unpublished data. Instituto de Investigaciones Clinicas y Medicas, Madrid.

Wegstein, J. H. 1958. Accelerating convergence of iterative processes. Communications of the Assoc. for Computing Machinery 1:9-13.

Willcox, O. W. 1959. Footnote to freedom from want, a quantitative appraisal of the food and population problem. J. Agr. and Food Chem. 7:813-822.

Wilson, D., Bressani, R. and Scrimshaw, N. S. 1961. Infection and nutritional status. I. The effect of chicken pox on nitrogen metabolism in children. Am. J. Clin. Nutrition 9:154-158.

Wissler, R. W., Steffee, C. H., Frazier, L. E., Woolridge, R. L., and Benditt, E. P. 1948. Studies in amino acid utilization. III. The role of the indispensable amino acids in maintenance of the adult albino rat. J. Nutrition 36:245-262.

Wolontis, V. M. 1956. A complete floating-decimal interpretive system for the IBM 650 magnetic drum calculator. IBM Corporation, Applied Science Division, Technical Newsletter No. 11.

GUZMAN, MIGUEL ANGEL. Study and Application of a Non-Linear Model for the Nutritional Evaluation of Proteins. (Under the direction of ROBERT JAMES MONROE.)

This dissertation is concerned with the development and application of a mathematical model which relates the protein concentration in the diet to the observed efficiencies. In developing the model protein induced responses were assumed to obey a law of diminishing returns with respect to the concentration of protein in the diet, while the rate of change of protein input into the system was assumed to remain constant with respect to the dietary protein concentration. Accordingly, observed efficiencies (z) can be expressed as

$$z = \frac{\eta - \xi e^{-\gamma p}}{\alpha + \beta p} \quad ,$$

where η is the average maximum attainable response for all individuals of a given species; ξ is a scaling factor which varies with η and γ; γ is an index of protein quality; α is a general correction term for the general adequacy of the diet; β is a measure of "effective" food consumption, proportional to the observed food consumption; and p is the protein concentration in the diet. Of these parameters α and γ seem to be independent of species, while η, ξ and β depend on the animal species used in the feeding trials.

This model permits the estimation of at least three important attributes of protein quality, and should allow a better understanding of the meaning of protein quality. In this connection, the joint

consideration of the parameters γ and η might prove valuable for the understanding of amino acid unbalance.

Three ancillary constants which are independent of the species used in the trials can be calculated from the estimates of the parameters. A simple graphic procedure may be used for estimating the optimum dietary protein concentration (p_o), the dietary protein concentration necessary for maximum efficiency. The estimate of p_o can then be applied to the model to obtain the value of the maximum efficiency (Z_m). The protein concentration at which neither gains nor losses occur (p_b) can be calculated from the simple relation

$$p_b \doteq \frac{1}{\gamma} \ln \frac{\xi}{\eta} \quad .$$

The three constants p_o, Z_m, and p_b also reflect protein quality, but additionally, because of their nature, should facilitate the introduction of economical consideration in relation to protein quality.

For the application of the model, the proteins tested must be fed over a range of protein concentrations. This procedure eliminates the usual discrepancies in results when the biological trials are conducted using different single protein concentrations. This feature of the model made somewhat difficult its practical application because experiments covering a sufficient number of dietary protein concentrations are not abundant. However, some data obtained with either mice or rats by different investigators, using either weight gains or carcass nitrogen gains as response criteria, were available in the literature and were used for fitting the model. The protein sources covered in these

trials included dried whole egg, skimmilk, casein, amino acid enriched corn gluten, heated soyflour, INCAPARINA and wheat gluten. The results from these fittings indicate that the proposed model for the nutritional evaluation of proteins gives useful information for a better understanding of the factors which determine protein quality. The results described, however, suggest that additional research is necessary for the improvement and better characterization of the system.

STUDY AND APPLICATION OF A NON-LINEAR MODEL FOR THE NUTRITIONAL EVALUATION OF PROTEINS

by

MIGUEL ANGEL GUZMAN

A thesis submitted to the Graduate Faculty
of North Carolina State College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

DEPARTMENT OF EXPERIMENTAL STATISTICS

RALEIGH

1961

APPROVED BY:

Chairman of Advisory Committee