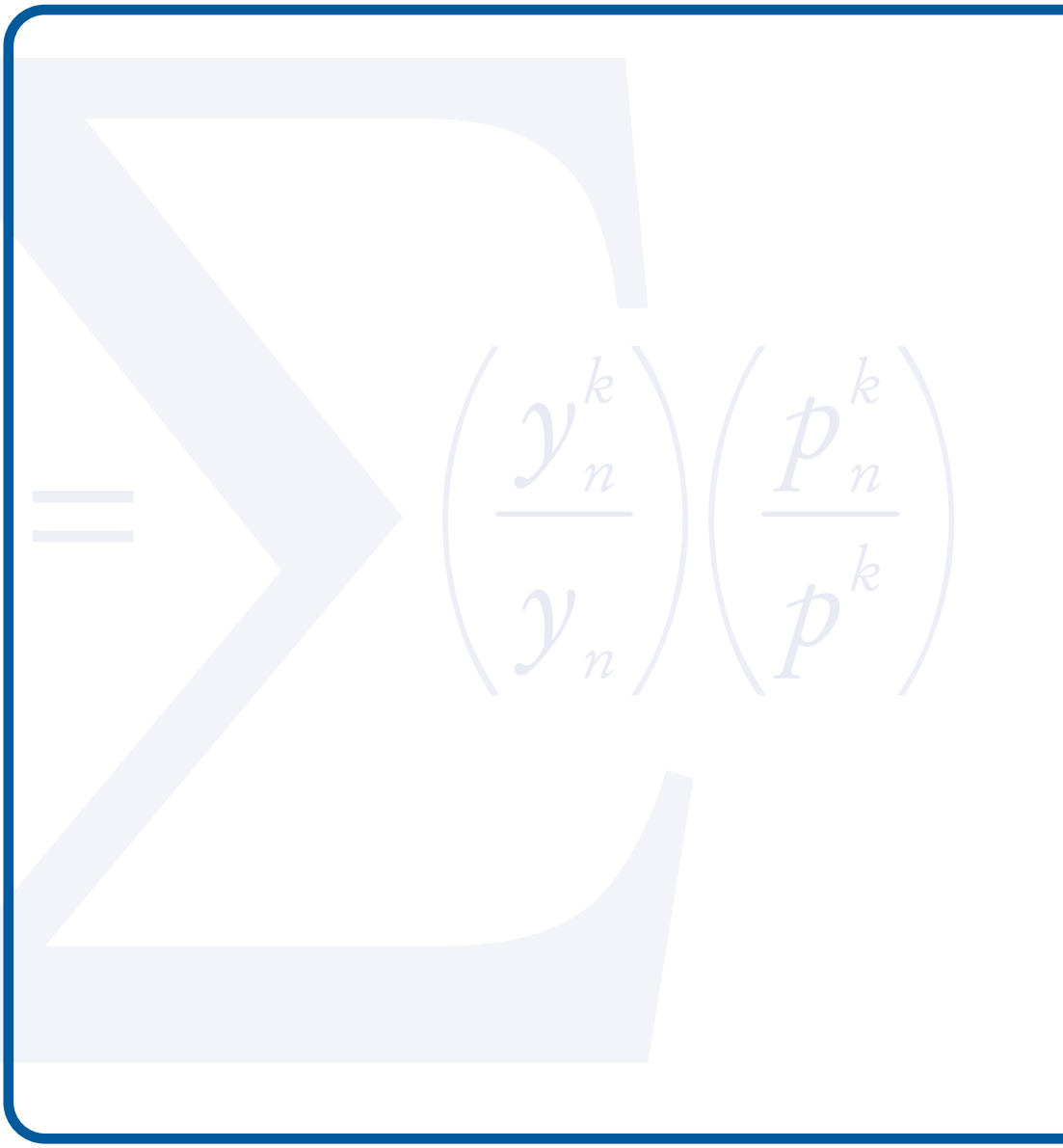
# Statistics Division

Working Paper Series

**NO:** ESS /




**FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS**

# THE PROBABILITY DISTRIBUTION FRAMEWORK FOR ESTIMATING THE PREVALENCE OF UNDERNOURISHMENT:

## *Exploding the Myth of the Bivariate Distribution*

Loganaden Naiken[1]

(logan.naiken@intnet.mu)

[1] A former Chief of the Statistical Analysis Service, Statistics Division , FAO

Page

*SUMMARY*

*In his pioneering study carried in the early 1960's, Sukhatme had formulated the estimate of the prevalence of undernourishment in a population within a bivariate distribution framework where dietary energy consumption (DEC) and dietary energy requirement (DER) are considered as random variables. The evaluation of the formula required the specification of the joint distribution of DEC and DER. In the absence of data on the joint distribution Sukhatme had, as an approximation, formulated the estimate within a univariate distribution framework involving the distribution of DEC and a cut-off point reflecting the lower limit of the distribution of DER. FAO's methodology for estimating the prevalence of undernourishment has been traditionally based on this univariate distribution framework. However, since this approach appeared to ignore the risk of undernourishment at DEC levels overlapping the range of variation of requirement, it has been criticised as yielding an underestimate of the magnitude of the problem of undernourishment. In view of this some analysts have attempted to apply the bivariate distribution framework by modeling the joint distribution of intake and requirement. Others have applied the univariate distribution framework but used the average DER requirement rather than the lower limit of the distribution of DER as the cut-off point. All these attempts have led to very high estimates of the prevalence of undernourishment. In further studies undertaken in the 1970's Sukhatme has attempted to justify the univariate distribution framework that he proposed earlier by postulating the theory of intra-individual variation in energy requirement which implies that an individual cannot be considered to be undernourished or overnourished as long as his or her DEC is within the range of variation of DER. Since the variation in DER has been traditionally considered to be a reflection of differences between individuals, i.e. inter-individual variation, the theory of intra-individual variation has instead led to controversy and dispute rather than an understanding of the basic principle justifying the validity of the univariate distribution framework. This paper reviews the debate surrounding the issue since Sukhatme's pioneering study in the early 1960's and points out that the primary source of the controversy and debate has been the failure to realise that the distribution of DER in fact represents the realization of the joint distribution of DEC and DER with the consequence that the probability of DEC being in balance with DER is high for the DECs overlapping the range of variation of DER. Thus it is in fact the latter that explains Sukhatme's argument that an individual cannot be considered to be undernourished or overnourished if the individual's intake is within the range of variation of requirement. The failure to realise this has led to the continued belief in the myth of the bivariate distribution framework and hence the application of flawed models of the joint distribution of DEC and DER.*

***Key words:*** *Distribution of dietary energy consumption (intake); distribution of dietary energy requirement; joint distribution of dietary energy consumption(intake) and requirement; bivariate distribution framework; univariate distribution framework; inter-individual variation in dietary energy requirement; intra-individual variation in dietary energy requirement; lower and upper limits of the distribution of dietary energy requirement; correlation between consumption and requirement; probability of consumption being below requirement; probability of consumption being in balance with requirement; probability of consumption being above requirement*.

# I. INTRODUCTION

FAO has been traditionally estimating the prevalence of undernourishment on the basis of food consumption data (expressed in terms of dietary energy) and dietary energy requirement for the purpose of quantifying the dimension of the food inadequacy problem particularly in the developing world. In this connection undernourishment has been defined as the state whereby dietary energy consumption (DEC)[2] is below dietary energy requirement (DER)[3] and the undernourished refers to the individuals in this state. The prevalence of undernourishment has been defined as the proportion of the undernourished in the population.

The measurement of undernourishment on the basis of food consumption data expressed in terms of dietary energy may be justified from two perspectives. Firstly, a certain amount of dietary energy is essential for the maintenance of body-weight and work performance. Secondly, an increased amount of dietary energy, if derived from normal staple foods, brings with it more protein and other nutrients as well, while raising the latter, without ensuring a certain amount of dietary energy, is unlikely to be of much benefit in terms of meeting food needs. In fact the amount of food consumed by individuals is best expressed in terms of dietary energy. Therefore the measure of undernourishment based on DEC differs from those based on anthropometric indices in the sense that while the former reflect food deprivation or hunger the latter reflect the broader concepts of undernutrition and malnutrition[4].

Another distinction between the FAO measure of undernourishment and the measures of undernutrition or malnutrition based on anthropometric indices concerns the unit of data collection. Disaggregated information pertaining to DEC is normally derived from the household food consumption data collected in national surveys and hence refer to households whereas anthropometric data pertaining to weight and height refer to individuals. Thus, while the measures of undernutrition and malnutrition are calculated on the basis of individual data, the FAO measure of undernourishment has to rely on household level data. Although the household data are normally expressed on individual basis through division by the number of household members or consumer units, the reference unit of the data remains the household and hence the intra-household distribution is not taken into account. In using the household level data for estimating the proportion of

---

[2] In nutritional literature it has been common practice to refer to intake rather than consumption of food or nutrients (energy, protein, fats, etc.) This is a reflection of the fact that nutritional status is determined by the food eaten or ingested by an individual (biological consumption) and nutritional requirements reflect the intakes of well nourished individuals. However, the term adopted in connection with the data collected in national household surveys and used for estimating the prevalence of undernourishment is actually food consumption In view of this the term "consumption" rather than "intake" is used in this paper except when quoting from the relevant nutritional literature.

[3] Dietary energy requirement refers to the human biological needs after taking into account age, sex, bodyweight and physical activity.

[4] It should be noted that in the past FAO had referred to its measure based on food consumption as "undernutrition". However, beginning with the *Sixth World Food Survey*, this practice has been discontinued in order to distinguish the FAO measure from the measures based on anthropometry that reflect not only food insufficiency but also adverse health and environmental factors.

individuals having DECs that are below their respective DERs it is assumed that food is distributed according to the needs of the individuals within the households so that if household DEC is equal to household DER, the requirements of all the individuals in the household would be met. In any case, the use of data pertaining to household access to food and attempting to capture the individuals in the households whose access to food are below their needs, the measure has the merit of referring to a basic aspect of poverty.

However, the food consumption data from national surveys refer to a probability sample of households rather than the totality of the households in a population. Moreover, the DER recommended by the international expert groups on nutritional requirements refers to an average for individuals classified by sex and age which means that the actual DER of an individual is not known. Thus the estimation of the undernourished in a population cannot be viewed as a simple accounting exercise involving the comparison of the observed household DECs with calculated household DERs and counting the individuals in the households that have been found to have DECs that are below DERs.[5] Instead it has to be viewed within a probability distribution framework where the estimate is actually the proportion of the population undernourished. The number of undernourished is subsequently derived by applying the estimated proportion to the total population. Work in this direction was initiated in FAO in the early 1960's through the pioneering study of Dr. P. V. Sukhatme, who was then Director of the FAO Statistics Division. This study, which was presented at the joint meeting of the *Royal Statistical Society* and the *Nutrition Societies of London* (Sukhatme, 1961), in fact laid the foundation to FAO's use of distribution analysis in preparing estimates of the prevalence of undernourishment that began with the *Third World Food Survey* (FAO, 1963).

Sukhatme had originally formulated the estimation of the proportion of a population undernourished within a bivariate distribution framework, where DEC and DER are considered as random variables. However, in the absence of data on the joint distribution of DEC and DER, he had formulated the estimate within a univariate distribution framework that involves the distribution of DEC and a cut-off point reflecting the lower limit of the distribution of DER. This univariate distribution framework has been used by FAO in connection with its periodic assessments of the prevalence of undernourishment. Changes or improvements have taken place over the years but these have mainly concerned the specification of the distribution of DEC and the calculation of the cut-off point.

However, Sukhatme's derivation of the univariate distribution framework has proved to be elusive or not convincing to many researchers. In view of this, some have attempted to apply the bivariate distribution framework by modeling the joint distribution of DEC and DER or the conditional distribution of DER given DEC. Others have applied the univariate distribution framework but used the mean rather than the lower limit of the range of variation of DER as the cut-off point. However, all these attempts have invariably yielded estimates that are too high to be realistic. Sukhatme later attempted to justify the univariate distribution framework by invoking the theory of intra-individual variation in DER which implies that an individual with DEC intake falling within the range of variation of DER

[5] Such an approach has been applied in a recent study by IFPRI researchers (Smith, Alderman and Aduayom, 2006)

cannot be considered to be either undernourished or overnourished. Therefore only the individuals with DEC falling below the lower limit of the distribution of DER can be considered as being undernourished.

Sukhatme's theory of intra-individual variation in DER had however aroused a major controversy and debate among nutritionists and economists and as a consequence proved to be more confusing than helpful in understanding the validity of the univariate distribution framework. Therefore the whole matter was reviewed in FAO in the course of the preparatory work for the *Sixth World Food Survey* (FAO, 1996). The review, which confirmed the validity of the univariate distribution framework, was discussed in a FAO staff article (Naiken, 1998). However, Svedberg (2003), claiming that the estimate formulated within the bivariate distribution framework reflects an "unbiased" estimate, had applied it by modeling the joint distribution in order to demonstrate that the FAO methodology and as well as data used for estimating the prevalence of undernourishment were flawed. In view of this, in an appendix of a subsequent paper presented by Naiken (2004) at the *International Scientific Symposium on the Measurement of Food Deprivation and Undernutrition* (ISSFDU) held in Rome in 2002, it was pointed out that the flaw was rather in the joint distribution models used by Svedberg and the others who have resorted to the bivariate distribution framework.

However, as indicated by Svedberg's comments in another paper following the Symposium (Svedberg, 2002), the argument was apparently still not convincing. This was partly because the paper presented at the Symposium failed to pinpoint the primary source of the flaw, which lies in the bivariate distribution framework itself. The aim of this paper is to highlight this point while discussing the history of the debate on the subject since Sukhatme's pioneering study in the early 1960's[6].

Thus, section **II** presents the probability distribution framework for the estimation of the prevalence of undernourishment and thus introducing the bivariate and univariate distribution frameworks as conceived by Sukhatme in the early 1960's. In section **III** the attempts made by Lörstad in the early 1970's to apply the bivariate distribution framework by modeling the joint distribution of DEC and DER are described. The FAO approach, which emerged following a methodological review in the early 1970's, is discussed in section **IV**. Section **V** discusses the approach of linking the measure of undernourishment