The Philosophy of Statistics and Probability applied in Soil Engineering

1. Introduction

The title was accepted as inferred from the plans of the Organizing Committee, but it is recognized right from the outset to be bold and premature, portraying much more a rationalized intent than a foreseeable achievement to be discussed in a State of the Art presentation. Truly, a more appropriate title would yet retain the separate steps of the ladder, and emphasize the intent of statistics, of probabilities, and of decision theory, as intended aids towards design in Soil Engineering.

Soil Mechanics can be said to have ow~~n~~ed its first important steps of success to cutting the Gordian knot from the complexities and vague qualifications of geology and the natural sciences of yore, and mentally assuming the arrogance of deterministic quantitativism, coupled with the unquestioned servitude to observations and tests. Mathematics is a deterministic idealization, and numbers, vectors, tensors and matrices are deterministic quantities connected by deterministic functions. It is very revealing to note that most soil mechanics publications, almost without exception, theorize in an absolutely deterministic cause-effect link, even in the face of very complex phenomena: on the other hand, tests and observations are a must, and almost without exception the results are presented and interpreted by curves joining successive individual points!

Thereupon, one must recognize that there is a very wide gap of frame of mind to be bridged first. If I may be permitted to draw upon a quotation from Prof. T.W.Lambe's recent Rankine Lecture (Geotechnique, June 1973 p.157) it may be seen that even where the spirit indeed is strongest (witness his intense enthusiasm for "Predictions in Soil Engineering") the conditioned flesh is weak: "The great variability of soil characteristics typically encountered by the engineer leads one to think that the principles of probability theory could be most helpful. In certain aspects of soil engineering, such as earthquake engineering, this expectation has proved to be true. In general, however, probability theory has not yet had a significant influence on the practice of soil engineering." Indeed, let us recognize that the final statement is a polite gross understatement; there has been no influence whatever. The first statement embodies what has become the unchallenged popular consensus "because of the great heterogeneities, the only reasonable manner to analyze data should be by statistics", and incorporates two further thoughts of great and complex pregnancy, that cannot so simply be bundled together with the mere notion of statistical variability.... that is, "Engineering", and "Probabilities". Finally, the intermediate statement exposes the ironical measure of the frailty of our degrees of belief: where there is absolutely no alternate (e.g. earthquake probability prognostication), no deterministic strut,

Proc. "Applications of Statistics and Probability in Soil and Structural Engineering".
2nd International Conference. Aachen, 1975. vol.III

and no possible measure of validity or success, the degree of returns of a creed are highest, and the belief is considered proven.

This presentation's* keynote will be Caution: as great as are the hopes, great can be the pitfalls.

We all know in our entrails the problems of heterogeneity, error, etc, that should be equated to permit transforming into palpable objective measures the "feel" for better engineering solutions that "experience" gives: that is, we are, we cannot help but be Bayesian in all our root culture! Yet through a curious psychological quirk of attempting to "deny the base steps by which thou didst ascend", the spirited vanguards armed with a new weapon want to prove they can start from scratch and give anybody, duly armed with numbers and formulae, the right to build the same castle!.

Actually, the moment that Soil Mechanics leaves its comfortable cradle deterministic "perfectly reproducible" links derived from controlled observation of idealized simple phenomena, it becomes necessary to recognize fundamental terms and concepts often misplaced:

Statistics fundamentally represents the method of reducing a group of measurements and data into a few values that retain the information contained in the original data: synthetic summary of data about samples, of events observed.[1] Statistical idealization is associated with a frequency, large number of repetitions, and the quantities used are defined by distribution functions or densities of probabilities.

Regression analysis comprises the attempt at statistically correlating parameters subject to independent random variations. When conducted with full use of statistical support, it not only permits crosschecking, but also serves for estimation and probabilistic extrapolation.

Probabilities apply to events that have happened, may have happened, may be happening, or may yet happen (universes that we imagine) and, being assumed to persist in the same trends of the past, are forecast to happen in the future. The very purpose of probabilities is connected with forecasting.

Strategic idealization (Decision Theory, Theory of Games) is associated with a degree of confirmation, dependent on a single action resultant on a decision, employing probability as a counterpart of a degree of confirmation and not only of a frequency.

Decision Theory is strategic, related to action (subjective or not) on the basis of probabilities, but depends on the value objective selected.

Prediction comprises the estimation of the most probable complex result from the interplay of probabilistic forecasting on several items. We shall limit the term probability estimation to a relatively small chain of interconnected events intervening, while prediction presently embodies a considerable amount of deterministic theorizing, and certain ambiguities regarding interdependence or not of varying components.

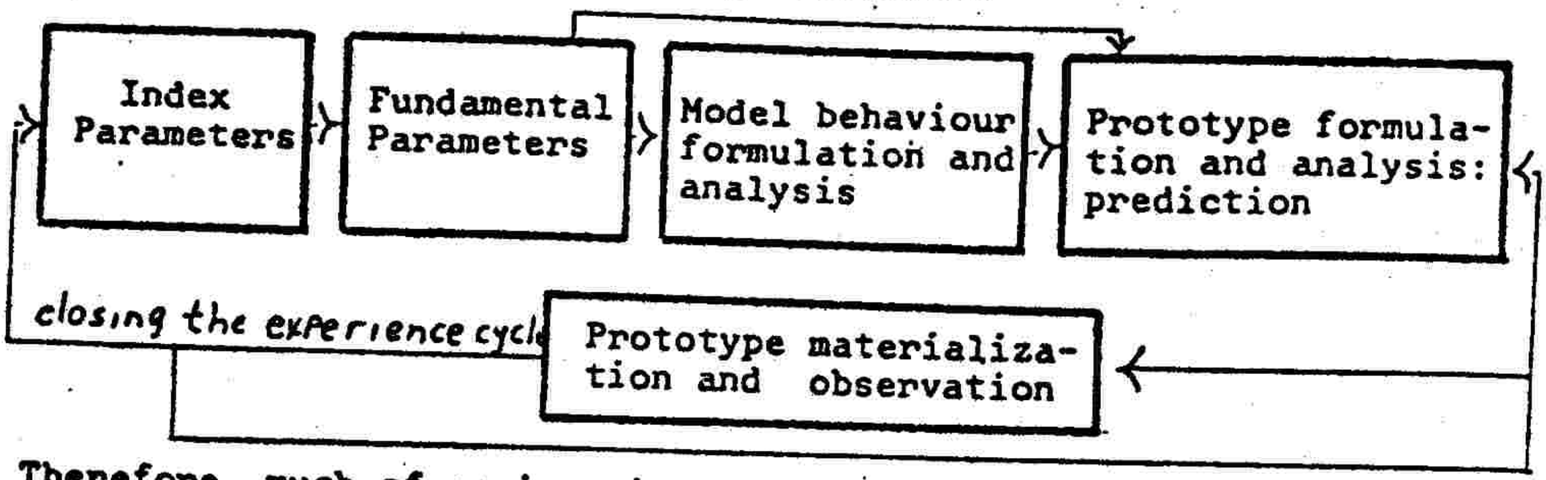
Prescriptions have historically and irrevocably oriented most of Soil Engineering Design (and, in fact, Civil Engineering Desing), and continue to do so. Besically, Engineering has advanced from a condition of application in affluence, wherein neither statistics nor probabilities nor decision strategy could have any place, to the gradually tightening requirements of optimization and of testing the frontiers of impunity.

It would be a delusion not to recognize that Soil Mechanics computations have followed in the wake of Design by Prescription, since one need but observes the curious adjustment of the technology to a great range of Factors of Safety and Levels of Acceptance it cannot really comprehend. In statistical terms a Prescription may be classed as "the choice of a change of universe for a solution, based on a freedom to place oneself safely outside of the universe under analysis and estimation". Thereupon, hitherto most successful engineering design has transcended Decision Theory's immediate limitation (limited, temporarily, to optimization of decisions within a given universe), through Prescriptions, and by using engineering as derived from ingenious (with an i) instead of as associated with engine, as well pointed out by Rosenblueth (   ). Prescriptions can improve with statistics, probabilities, strategy, and decision theory, and, above all represent the wisdom to shift to another universe of problem, sought to bypass the frustating uncertainties of the earlier formulation's universe.

Design really implies a decision of transcending that of decision theory as such, because of Prescription, and because of the ability to act deterministically on the built structure in such a way as to make it the overriding conditioner. Possibly the most commonplace foundation engineering example can be said to be the use of piles to go through the questionable layers of subsoil: and perhaps it is not without significance that in such piling design, "factors of safety" have been wilfully high. Statistically and probabilistically the change of universe and change of value objective does not contitute withdrawal of the design decision from analysis and judgement by decision theory. However, while the weapons are in the forging I would merely guard against the admonition "fools rush in where angels fear to tread". Good design is not yet cornered, from its position of affluence of ingenious ideas, into being better calculation or better estimation of intangible risks.

The time of more pressing demands will come. Let us use profitably the intervening time to build up a sound ladder.

As a final introductory thought I feel bound to emphasize two important facts that condition Soil Engineering very strongly. First, the fact that no matter what the external appearances may belie, the action sequence always involves.



Therefore, much of engineering decision is hidden behind the intangibles conveyed by index and fundamental parameters that have been deterministically invented and represented in an oversimplified manner: as we formulate the very basis for our engineering problem, we depend on a complex experience factor that is represented by the crudest tests.

Second, as an Engineer I always know a priori that any parameter X is a function of a great many parameters, X = f (a,b,c....z) and therefore, it is only through deterministic simplification that the bases of Soil Mechanics were laid on presumably valid index parameters and simplified "fundamental" parameters, all of which can be shown (as many have gradually been) to be insufficient in representation of the universe. The hope had been that the "series of intervening parameters" could prove rapidly convergent, so that the first, or first and second, or at most first-second-third parameters, could so closely define the universe, that the error or uncertainty from neglecting the remaining n parameters could result insignificant in comparison with other inexorable errors in the design process. Unfortunately, however, not every parameter (index or otherwise) could have chanced to meet the requirement of being the overriding significant parameter: and yet the closed-cycle deterministic theorization, and design prescription, has been built around it!

Where does one start revision, with estimation of optimized returns? That is the question: and, whatever, the answer, the admonition is CAUTION.

2. A cursory discussion of some of the partial challenges and present status.

Probably the principal difficulty presently faced in the application of statistics, probabilities, and decision strategy to Soil Engineering derives from the lack of significant hard data with respect to distribution, reproducibility, standard deviations and coefficients of variation, separation of "inherent" and statistical uncertainty (the latter reducible by increasing number of samples), etc. On the basis of a reference survey into some of the principal sources of geotechnical publications the following summary of information and comments is offered.

2.1. It may be noted that a few publications which report data in a manner x ± Δx, recognizing an observed range, are discarded as deterministic. Moreover, a number of publications have reported data in the form of histograms: apart from the usefulness in indicating the general configuration of distribution functions, these papers are also set aside as not quite belonging to the group belonging to statistical treatment. Quite frequently, as a matter of fact, the representation of test data through histograms of separate parameters has hindered the desired appreciation of the available information since it hides interrelationships (for instance, a clay layer described by side-by-side histograms of the Liquid Limit and Plastic Limit, since it is not recalled ever having encountered a histogram such as of a ratio LL/PL). (39, 51, 144, 172, 195). Soil Mechanics properties even though rather arbitrary (and possibly random functions at the Soil Science level) are more or less causally interrelated.

There has been a very dominant deterministic trend towards single-parameter correlations, and, in reflection of the eminently experimental bases of the technology, a very strong preference for graphical presentations; moreover, as has been jokingly said, the civil engineer has always found a way to linearize all trends, be it in cartesian coordinates, or in semilog, or if everything fails, in log-log scales. As a matter of side interest mention may therefore be made of the scattered instances of graphical presentation of histograms attempting to relate two histograms (146, Fig.2.1) attempting to relate two parameters, and so also, in triangular diagrams (146) or in cartesian space representations (165) the configuration of three.

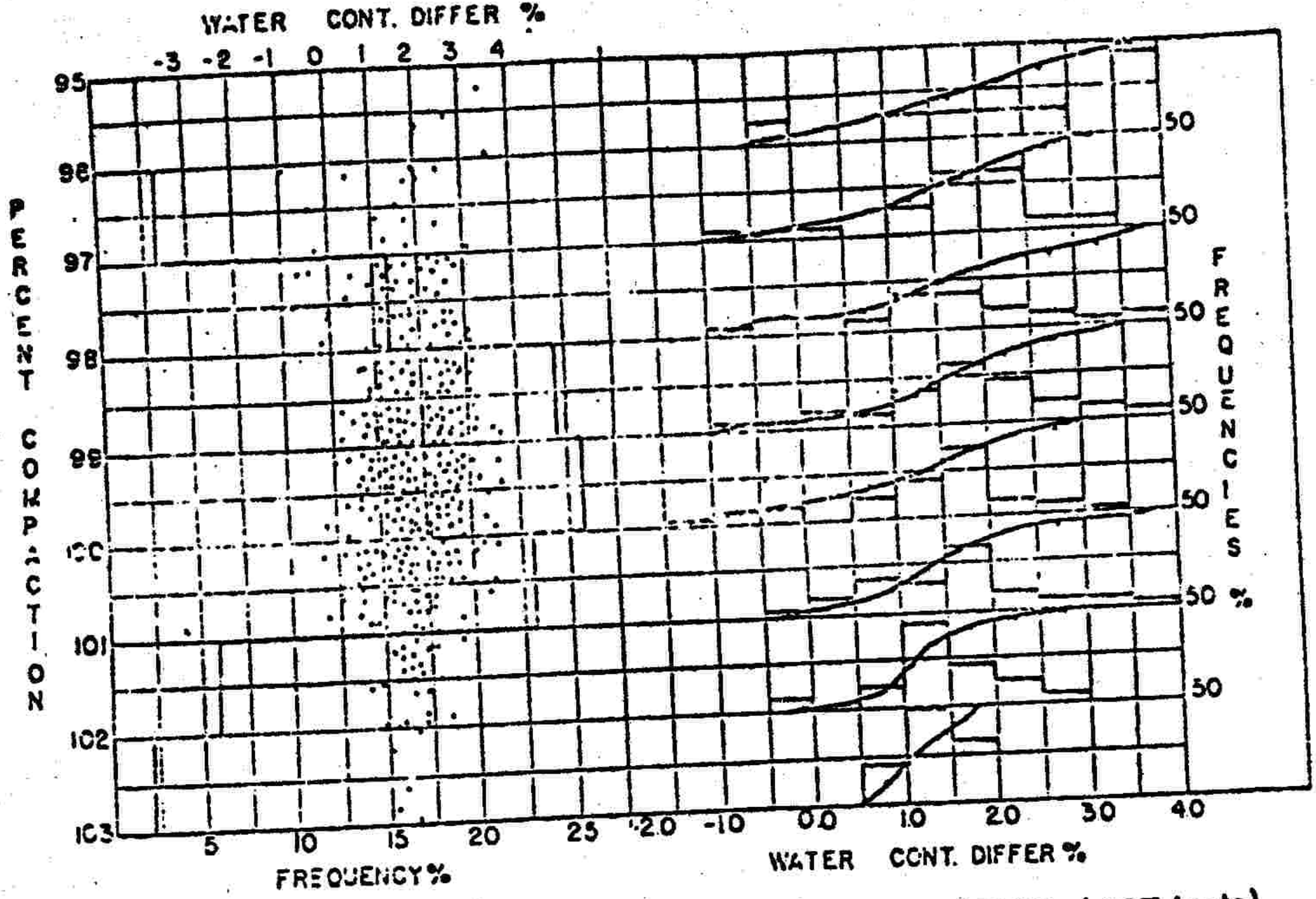


FIG. 2.1 SALTO OSÓRIO COMPACTION STATISTICS (467 tests)

2.2 Despite the seduction of graphical presentations, it must be recognized that within the realm of statistical representation, in Soil Mechanics and Soil Engineering we must accept nothing short of multiple regressions. It has been emphasized that in engineering reasoning one must a priori comprehend that infinite are the parameters that can and may influence, and therefore, the multiple regression should be the starting point, only reducing the number of parameters inasfar as their influence can be demonstrated secondary. In the context of engineering action (analogous to surgery as contrasted against biomedicine) it has already been emphasized that, consciously or not, one must seek overriding dominance of a parameter.

Some points of interest in connection with these principles may be summarized:-

a. Most regressions hitherto have been between two parameters; there are a few instances of double regressions (37,165) and even fewer of triple regressions (6). J.Neil Kay's paper presented at this conference reports on the first instance, to the senior Reporter's knowledge, of a multiple regression of more than three (as many as 13) parameters, and it may be of significance to ponder on the fact that the computer program was from the biomedical field.

b. It is regretted that despite statistical evidence and "theoretical justification" of the importance of a second parameter in a certain regression, many a subsequent effort takes no note of the fact. For instance, it was proven(23) that virgin compression index Cc must correlate not merely with type of clay (through LL) but also with in situ condition (through void ratio, e, for instance): this latter parameter can be reasoned to be most important in secularly consolidated clays, because of secondary compression quasi-preconsolidation; unfortunately not a single correlation of Cc=f(LL) subsequent to 1962 has remembered and/or investigated the interference of e.

c. Doubtless one of the principal reasons for the present frustrations in statistical predictions, and a fortiori in the more ambitious decision theory applications, lies in the fact that the "experienced soil engineer" always insists on a"complete" cross-section test program, even if only with very few tests of each type, in order to fully comprehend the "personality" of the soil (.61). There is always at the back of the problem the dualistic differentiation between a continuum of reality and the discontinuity of decision.

Thereupon, it must be recognized that most tests and computational routines were made to be overriding, deciding: and in moving from the statistical fact to the probability and decision world, the intervening parameters were deterministically forced to be discriminatory, dominant. For instance, the oedometer test has the trappings of a model test to represent the deep clay layer sandwiched between two sand layers, being subjected to unidimensional compression. But, early Soil Mechanics decided (and wisely) to apply "instantaneous, big" pressure increments, setting aside the one or two investigations on slow lead-shot loading, etc.; the pressure increment ratio of 2 permitted developing an arbitrary test conditioned to a principal overriding loading experience. Moreover, in dualistically subordinating clays into normally consolidated vs. preconsolidated, a graphical procedure was developed for sharpening the differentiation (truly rather smoothly transitioning in nature) and thereupon in most computations the absolutely unnecessary assumption (today unperceived most often) is made, that settlements are zero up to be p_c value and most significant at immediately higher pressures (because of the ratio of logs). Was it not right engineering decision? But is such arbitrariness conducive to appropriate statistics? In the same light was not the creation of the rough SPT boring penetration index a valid first-order engineering decision approach? And the use of limiting condition routine tests, fully drained vs. absolutely constant volume undrained, for a basis for prescriptions? And the definition of critical void ratio, not for forecasting under what conditions a sand mass may liquefy, but for prescribing under what conditions it will not liquefy? And so on?

In short, it appears important, to the senior Reporter, to recognize that most of the parameters created and in use consciously or subconsciously were directed towards being dominant and not necessarily favouring appropriate description of a natural, non-dominant multifaceted reality. In such a situation it appears that for optimized feasibility of the desired applications of probability, a two-pronged attack must be pressed: on the one hand, highest possible use of Bayesian support, even though temporarily employing existing stilts, and guarding, by "experience" against pitfalls of preconceptions; on the other hand, gradual revision of the very tests and parameters inasfar as convenient..

2.3 Within the use of regressions, several problems arise, that have been noted and mentioned but which may yet profit from some attention.

a. There is the choice of what parameters to concentrate on such that they may yield correlations reasonably plausible (37) not merely as regards the statistics, but also as regards insertion into available theorization (deterministic) and experience (Bayesian reflection). For instance, the use of Consistency Index IC was rapidly dropped in view of the demonstrated influence of structure and sensitivity of clays. But there are still many instances of attempts at correlating Cc vs. water content W (26) or allowable bearing pressure on footings conditioned by settlements vs. SPT. It must be remembered that the geotechnician has "created" many parameters and indices, not so much because they were "proven" (a fortiori under the publish or perish pressure) but because he wanted to throw them into the arena for a survival test.

b. A second question concerns the investigation of "functions" of a given soil parameter of interest, in order to improve the tendency towards developing a Normal Distribution, for instance (146). On the one hand, it is reported that considerable advances have been made in the uses of distribution-fee statistical tests. On the other hand, regarding the physics behind measured and computed parameters, it may be worthwhile examining each case to check if no compensating errors or arbitrary algebraic consequences are being unwittingly, introduced. For instance, in quality control of embankment compaction even though it was recognized that the more significant parameter is the water content difference ΔW from the optimum, for many years no other means were available but to determine this parameter by subtracting one test value from another test value: obviously the Hilf technique of direct determination represented a very significant improvement in principle and practice.

c. Another important question concerns striking a wise balance between fruitful Bayesian prior conditioning of search for statistical data and correlations on the basis of assumed theorizing, and allowing enough flexibility so that the correlations may also iteratively revise theorizing. We must conceptually recognize that much of the very application of statistics has been deterministically straightjacketed into pseudo-theoretical expectations, such as the presumed linear correlation between SPT and cone point resistance Rp (37) At the other extreme, when we do fundamentally depend on a theory (such as Mohr's theory and the Mohr-Coulomb straight--line representation of the strength equation), within the pragmatic arbitrariness of the notion of fertility of a particular idea and practice how can we accept a statistical procedure for regression (9) which in its derivation implies a strength criterion related to maximum shear stress (or principal difference) rather than maximum stress obliquity? It may be noted that for a very simple case a straight line regression based on points corresponding to maximum obliquities has been used (34 .), and for long-term cases corresponding to preconsolidated or compacted materials the complete envelope has been subdivided into three universes for three straight--line regressions: above pc a normally consolidated linear regression through the origin; for over consolidation ratios OCR of the order of 1 to 8 a straight line regression with c and ϕ ; and finally, by decision or prescription, for high OCR values wherein shear may cause negative pore pressures, prudently assumed untenable except for short-term conditions, a final regression forced through the origin once again.

d. Yet another important problem concerns the interpretation of the physical reality and applicability of extreme values, how to fit in missing values, and what to do about truncated histograms. There are, in many an instance, quite defendable "physical realizations" that there are absolute minima or maxima, not only of properties, but of volumes of material within which the property has to be "averaged" (volume assumed homogenous) in order to begin to be significant within a bigger mass or problem: so one wonders at an assumption, for instance, in a sedimentary clay stratum, that "the lower bound of strength is zero" (120). On the other hand the question has been raised, for instance (36) as to the difference, in a hard fissured clay, between the universes of shear strength determinations through laboratory specimens and through in situ tests, when in the specimens one does not take into account the observations of so low a strenght that the specimen failed on trimming.

How big a volume of a sand should, at the extreme of the distribution curve, be necessary to undergo erosion of liquefaction in such a manner as to trigger and spread the phenomenon? Considering that minimum tensile strength at a small soil element may be zero, how big a volume of soil element would have to be considered for tensile failure in order to develop potential mass instability? How would histograms, and particularly extreme values thereof, change as the "minimal consequent volumes" change?

2.4 Some of the more immediate tasks of the steps to the desired overall ladder may be pointed out.

a. One neéds to challenge some of the early parameters, particularly the Index Parameters (38) and the parameters wilfully introduced (associated with corresponding tests) to be dominant. One needs to invent new parameters, in order to be more suitable for probabilistic decision calculations: such proposals as the improved quantitative representation of soil grading (103)or the new plastic ratio chart(143) are interesting examples. Some of the early parameters apparently could be conceptually discarded right from the start (47): for instance, the Relative Density RD parameter appeared to have been created under a valid concept, similar to the IC, but with one important difference, that it would depend on the difficulties of defining and measuring extreme values. In general, for the creation of more appropriate parameters one should attack the constitutive laws (see item 4 below) for orientation, and, as mentioned in 2.4 (c) below, the first useful statistical confirmations should be obtained in extrapolation of the very test curves available, defining the law through the first n points and forecasting the stretch from n+1 to n+m. It must be emphasized that most investigations and preocupations have been with respect to precisions, accuracies and reabilities of soil parameters (e.g. 160); it appears, however, that one of the questions that must be challenged is the significance, and discriminanting significance, of the parameter within the problem.

b. Not enough attention has been dedicated to statistical orientation and to program delineation for testing, both research activity and in professional work. Curiously some work along this line has advanced more in the complexities of field situations (8, 64) than in laboratories from which the "fundamental laws" continue to emanate deterministically. One reaches surprising conclusions when one runs ten would-be identical tests and even when one applies such information on a would-be absolute rejection criterion (33).

c. Both in Soil Mechanics and, subsenquently, in Soil Engineering some problems of simple extrapolation shoud be statistically defined and verified first, as parts of the overall ladder. One such concerns the well-known problem of model-to-prototype laws of similitude (46): as is well-known, depending on the assumptions of brittle rupture, ductile rupture, or failure by deformation, the increases in dimensions affect stability in a different manner; should not the question be investigated first within Soil Mechanics rather than advancing directly into postulations of "progressive failure" etc. in complex Soil Engineering situations?

There are obvious physical limitations as to how big or how small a volume can be considered,and adopted as the basis for reasonings, irrespective of the fact that in highway engineering one may not be interested in minor variations, and in piping erosion one may be interested in trigger reactions.

Moreover, as has been pointed out employing such a study as reported by Bishop et al. (1973, 8th ISMFE., Moscow, 1.1p. 57) it would be of utmost interst to establish statistical lews for a stretch of the stress-strain curve, and extrapolate, much as was done for A and B coefficients without the prior assumption of being constants (165) and in a similar line has been done by I.M.Smith" Incremental numerical solution of a simple deformation problem in soil mechanics" 1970 Geotechnique 20,4, p.357.

2.5. Finally some comments on considerations connected with Soil Engineering, and design.
For the sake of clarity it may be stated that an arbitrary pragmatic distinction is beingmade herein between Soil Mechanics theory, derived and proven with reference to laboratory tests and very well controlled model tests and small-scale idealized field tests, and Soil Engineering theory with reference to field observations, generally very strongly conditioned both by former, and by a posteriori deterministic analyses that are fitted to give the desired result.
Design is concernend with the latter, notwithstanding the strong influence that the former exerts on all quantification, analysis and synthesis. Some considerations that come to mind may be summarized.

a. It has been assumed that the geometry may be treated as subject to very small error. That depends, however, very strongly, on the discriminating power of the classification used to define the geometrical boundaries and the presumed behavioral discrimination that the theory imposes. In short, it must be recalled that practically all the theorization (deterministic) of Soil Mechanics is based on the subdivision into the idealized cases of "pure clays" and "pure sands", and routinely relyes on quite different treatments for each case (for instance, settlement problems in clays based on consolidation tests and theory, while similar problems in sands based on the Buisman method or on plate tests).

b. Throughout the line one finds the assumption that in soil engineering design one is faced with the problem of very limited data (and thereupon, the great advantage of the formal tool of Bayes' theorem, the theorem of the probabilities of the hypotheses).

It seems that the problem is much more of a considerable storage of indirectly conditioning data that, in the face of a given professional problem may not be felt and brought to the fore in a first cycling of the computations, but may iteratively suggest a revision of the very parameters and probabilities earlier assigned. For instance, the condition of an upper dried and preconsolidated crust in clay deposits is generally accepted, and, as a first hypothesis one would assume normally consolidated conditions within the clay stratum: however, upon closer thought of the formation processes one can and should frequently postulate drying conditions at lower levels too. In other words, "experience" may not act merely on the input parameters but most frequently, on the output. A soil engineer with experience on buildings in Santos may not have as much analytic ability to assign appropriate ranges for parameters, as the may have to compare the final computed settlement and to "know" that a certain result is plausible or not: one must only guard against wishful thinking. Regarding reliability, the probability of a probability, frequently experience intuitions are better and easier in connection with the complex lumped-parameter end result.

Considering the step-by-step manner in which any project moves, one may always assume a known condition on which to proceed, and the good design is one that minimizes the types of change and the magnitude of change at the important operational phase. As regards factors of safety, it is interesting to note that quite absurdly no distinctions are made between projects that are subjected to great and sudden changes of stresses, and those that are not: for instance, in foundation design the problems of failure are totally different(and case histories so confirm)in tanks and silos of slow dead load application of about 10% of the total and very rapid operational loading of the remaining 90%, in comparison with reinforced concrete highrise buildings applying a dead load of about 85% over a construction period of 18 to 24 months with the final live load increment of only 15%. It is in a sense Bayesian to base engineering design not so much on computations of factors of safety FS as on minimizing changes of factors of safety ΔFS associated with the conditions imposing greater risk (38). For instance, as a fill or cut slope grows to the greater final design height we possess prior "knowledge" of the stability at earlier stages; and it may be of considerable significance to note that the parametric cohesion intercept and $ds/d\tau$ slope, change with increasing strains and changing obliquities, and especially, the proportional contribution of the more erratic c component of resisting forces dwindles in significance.

Proverbially there is the last straw that breaks the camel's back---which presupposes a brittle failure or a very well defined final level of acceptance or value system . In reality, however, there are very many straws that make gradually more unbearable to the camel, which merely means that we are part of the very value systems that we set up, and therefore they stretch or tighten as conditions change. Which is really one of the hidden reasons why the mathematics of probabilistic decision-making is still too deterministic for the value systems of life, of the dynamics of life, of life's search for itself.

d. Regarding the systematic establishment of the merits of the newly proposed computational approaches, it is often thought that one has to await too long a process of monitoring new projects in order to be able to develop the checks on the probabilistic theorization. Not so: all that one has to do is to use past cases (possibly aided by some fitting-in of data) with concious guard against wish-conditioning. Moreover, in engineering cases, rather than merely analysing a new case in the light of supposed data from n other cases, one could gain much information from analyzing the very developments of conditions within the one case: for instance, in a building undergoing differential settlement, knowing the first n points of measurements, forecast the following one or two years, or knowing when the cracks started along some wall panel, attempt to establish the probabilities of new cracks developing, and so on.

Partial problems such as those above mentioned may not prove too attractive, but will be necessary.It is quite understandable that in the face of each complex engineering problem the desire is to build up directly the complete decision tree, to achieve results that may prove more attractive to the profession.