*Proc. 6th Conference on Soil Mech. and Found. Eng.,
Budapest, 1984 (G. Petrasovits ed.)*

DESIGN DECISIONS, DESIGN CALCULATIONS, AND BEHAVIOR PREDICTION COMPUTATIONS: REFERRED TO STATISTICS AND PROBABILITIES

DE MELLO, V.F.B.

Rua Frederico Chopin 190,
01454 Sao Paulo, SF

SUMMARY

The insertion of concepts of statistics and probabilities into geotechnical engineering design continues to provoke intense debate. If we examine the history of Science, and the furious battles that raged between any novel theory and the entrenched beliefs, we well understand the present clash. Of course, everybody can see that the use of determinism and nominal Factors of Safety FS should rapidly decrease, because there are dispersions /statistical/ and these should condition probabilities of behaviors different from the one computed. But Civil Engineering does not deal with multiples as do all industries and no two prototypes are identical, no models truly represent the prototypes, and, above all, in big civil works of great responsibility we deeply recognize that we cannot permit, cannot conceive of permitting, catastrophic failure /thus seemingly implying that determinism is impossible to uproot/. How can we conciliate the two schools of thought, that appear totally divergent?

In this brief essay I shall attempt to explain my views which at least offer some hope of idea-fertility, and open perspectives of important advances in the collection of digestable data from prototypes.

STATISTICS OF EXTREMES. CATASTROPHIC FAILURES AS SINGULAR EVENTS, EXTREMES

In the statistical theory regulating the variabilities of a presumed fixed physical universe, which respectfully obeys given mathematical postulations, there have been some brilliant deductions of extreme-value probabilities. Some of these have found considerable use /e.g. in some versions of spillway hydrology/. In my Rankine Lecture /1977/ I showed that even in the small number of most-used extreme-value distributions, at the desired really low probabilities, the computed results are really very different. Moreover, in several cases of de facto test observations of great numbers of tests /e.g. industrial multiples etc./ it has been found that the probability functions do not fit the sufficiently rare events. Thus I find myself fortified in the conviction that catastrophic failures of civil geotechnical works are not amenable to computational analyses. Each civil work is somewhat singular /especially in natural conditions, geotechnical etc./ and catastrophic events are /fortunately/ rare in this random universe /e.g. the only thing common in "dams" is the name-designation that we give them/; one cannot establish anything better than spurious statistics in such a condition /e.g. "most persons die in bed, therefore bed is the single most dangerous place for humans"/.

The important saving grace is that generally we learn much from a failure, we perceive the physical mode of failure, and thereby tend to solve the extreme-value problem by a choice of change of physical universe, by a change of type of structure. For instance, the frustrating events of catastrophic floods /and some consequent dam failures because of insufficient spillway capacity/ are leading to the awareness that the probable maximum flood PMF, or 1:10 000 year recurrence flood, for design, are not fixed values at a given site: after a first estimate of the desired spillway discharge capacity, one must revise this very estimate of peak incoming flow /basic design data/ by reference to the flood volume, rate of rise of reservoir with incoming flood, ratio of surcharge to flood volumes, and "rate of consequence" to the dam in the remote case of overtopping; that is, the number porbabilistically computed is drastically adjusted to physical factors.
Thus, in the case of earth dams, the localized phenomenon of piping /extreme-value case, because it can start at a point and degenerate to catastrophe/ is not a type of phenomenon for which calculations of "factor of safety" have any meaning or place: the exclusion of the eventual condition is by an inventive physical change of statistical universe, e.g. by employing a fully-intercepting filter-drainage "curtain".

ENGINEERING CREATIVITY FOR AVERTING CATASTROPHIC FAILURES

One generally finds that a really inventive physical solution to avert a seriously problematic behavior, tends to be superabundant. The reasoning may suggest a sophism, since if it were not truly remarkable and noted, it would not be called inventive. At any rate, it is a very educative exercise to reflect on how much of engineering progress came by jumps, by the introduction of a physical change of statistical universe by inventiveness /Asian Conf. SMFE, Haifa 1983/. We act and do first, and then we attempt to compute around what was done and proved clearly successful: we would like to advance gradually tip-toe to the new frontier of impunity established by the invention. In dam design the use of grouting of cracked rocks, and the use of filters fully intercepting the flownet were physical inventions against conditions subject to extreme-value statistics; in tunnels and deep foundations, the uses of compressed air and bentonite slurry stabilizations were analogous; and so in reinforced earth, both Babylonian and modern, the imparting of tensile strength to a non-tension material.

In short, we should stimulate to the utmost the principal design weapon, which is ingenious engineering, the bypassing of a problem or setting it aside by use of new materials, procedures, combinations.

STATISTICS OF NON-FAILURE BEHAVIORS. HISTOGRAMS AND THEIR TRUCATION BY YES-NO DECISION OF DESIGN

Once an invented physical model has proven satisfactory, superabundantly so, our intent for the sake of progress and economy should be to observe a great number of non-failure conditions, at different degrees of proximity to the "failure". The failure itself is of journalistic interest, and serves to humble us,but does not arm us for design computations. We need to collect a vast amount of information on pre-failure misbehaviors in order to establish bases of statistical computations.

Many factors join in making this endeavour difficult. Firstly, the inventors make every effort to overdesign with their invention, both for commercial interest /royalties, patents/ and because while the behavior and possible failure is not yet understood, there is a desire and need to play extra safe. Secondly, unless and until failures occur, the mode of failure may not be clear, therefore the pre-failure misbehaviors become more difficult to program. Thirdly, precedent, standards, and inertia set in, "consecrating" practices insufficiently understood. Fourthly, there is no glamour in collecting such non-journalistic case history data. Finally, very few indeed are the designers and clients willing to try out on the same project /same statistical universe/ a varying degree of treatment expressly for the purpose of constructive curiosity: for instance, are there any cases in which, on a very long embankment, different stretches have been tried at slopes of IV:1.5H, 1:1.75, 1:2, 1:2.25, for the purpose of collecting information on Satisfaction Indices /Rankine Lecture 1977/ for adjustment of slope design to nominal FS values? In my view, this is presently the greatest obstacle to progress, and to compatibilizing the older deterministic designers and the applied-statistician decision-theory designers.

Such histograms inevitably have their statistical dispersions and percent confidence bands. Temporarily, however, let us reason as a single-line histogram /Seminar; Past, Present and Future in Soil Mechanics, M.I.T., Sept.1981/, and recognize that we can only advance if we completely separate the efforts at establishing misbehavior histograms, from the design decision yes-no, accept-reject, allowable-unallowable /which is much more conditioned by different societies/. For instance, in establishing unallowable settlements of buildings with regard to cracking, why not accept modest numbers and widths of cracks vs. discussing the intangible /extreme-value/ appearance of a first crack? And, for obtaining more interpretable histograms /one for each building, distinct physical universe/ why not gather varying data, floor by floor, obtaining varying differential-settlement data at bases of columns, one for each floor: by considerable simplification, to exclude secondary-stress sturctural effects, purely as regards settlement the top of the nth floor acts as if it were a foundation for the /n+1/th floor.

DESIGN COMPUTATIONS: ANCHIEVING PREDICTABILITY OF WHAT SHOULD NOT HAPPEN TO THE DESIGNED STRUCTURE

I have repeatedly emphasized that geotechnical design is not based on predicting what will be the behavior /Lambe/ but on what will not be the behavior /undesirable/. In other words, we design for limit hypotheses. Often for the same structure in two separate computations we assume first the maximum probable hypothesis, and next the minimum probable /opposite/ assumption. For instance, for an earth dam founded on a fractured rock foundation, with regard to flownet behavior across the dam we assume the foundation absolutely impervious /pushing the flownet upwards/, but with regard to foundation seepage or eventual piping into the foundation, we check under the hypothesis of a pervious foundation.

In statistical terms, we try to estimate the maximized width of the percent confidence band around our reasonable /average/ hypothesis, and we carry out design computations to try to assure that our project can satisfy conditions above or below /equal to or better than/ such confidence bands. We generally do not aim for a foundation design to achieve a settlement of 2±0.5 cm, but try to assure that the settlement will not exceed 5 cm. Decisions change by incremental quanta, and not as a continuum: e.g. for settlements up to 5 cm we may accept some solution, for settlements between 5 and 12 cm another, and so on.

Of course, once again, we must improve our establishment of the average /most probable behavior/, but principally we must narrow our confidence bands, because for economy /without impairing safety/ we want our design computations as near as possible, but just outside of, the upper or lower limits /confidence bands/. Assumptions of "instantaneous" reservoir drawdown, "fully saturated" embankment, etc., are so extreme as to demonstrate clearly our very low level of knowledge and confidence on the subject. Limit analyses and nominal Factors of Safety have to be understood as inevitable steps of a relatively ignorant past. A fixed FS number is absurd since it should depend on the dispersions and confidence bands involved. For improving our coefficients of adjustment of predicted behavior vs. observed prototype behavior, we are most often involved with prediction computations using the most probable parameters /statistical average regressions/. This is the important place for Lambe's prediction problem.

FACTOR OF GUARANTEE FG AND FACTOR OF INSURANCE FI AS COMPARED WITH FACTOR OF SAFETY FS

I have proposed /Symposium on Dam Engineering, Bangkok, Dec. 1980/ that whereas we have used only the concept of FS /ratio of Resistances ± errors, over Stresses ± errors/, even if we restrict ourselves to considering only dispersions around Resistances, we must recognize in our works the existence of two other Factors, FG and FI, which are quite different, especially when the dispersion is wider, so that FI FS FG. In order to clarify the concepts it may be convenient to exemplify with regard to piles, with which familiarity is greatest, and softground tunnels, in which "execution effects" are of greatest moment.

A pile jacked down under 60 tons to absolute stoppage of penetration/settlement has FG = 2 if used for a working load of 30 tons: by some lower rejection criterion /stoppage/, I have assured myself that the histogram of Resistances can only be higher than the value of 60 tons already pretested or guaranteed. Meanwhile, if the design estimated Resistance is 60 tons, the pile of working load 30 tons has the conventional FS = 2. Setting aside the discussions on dynamic vs. static resistances of piles and cases of sensitive clays, driven piles checked by "refusal" observations can well be said to imply factors FG. In contrast, a bored pile would suffer from two disadvantages in its load-settlement behaviour. Firstly, it would never have been pretested, and therefore one might conclude that it is affected by FS /poorer than FG/. Secondly, upon closer examination we should reason that it is even worse than that. All efforts of advancement of Soil Mechanics are towards minimizing sampling and testing disturbances, and better representing in situ soil parameters /intact soil elements/. In reality the assessed intact parameters would establish an upper rejection criterion, since the soil affecting bored-pile behaviour represents a histogram of resistances always lower, to varying degrees, truncated at the upper value. A situation diametrically opposite to that of FG, with the lower rejection criterion. One could denominate the new ratio of averages /Resistances/Stresses/ a Factor of Insurance FI: insurance is against something essentially inevitable, that should be attenuated. If projects continue to be designed generally for /nominal/ FS =1.5, without recognition of the significant difference FG FS FI, all structures in which FI is at stake will record a much greater degree of troubles, while structures in which FG is at stake will incorporate an unnecessarily higher degree of safety. Tunnels and bored piles involve execution effects that only deteriorate in situ parameters /resistance, deformation/ and therefore involve FI conditions.

In the case of dams and the eventual highly catastrophic downstream slope failure due to full reservoir, I have emphasized that the deterministic school insists that there should be zero probability of failure. The statisticians would argue that there is no zero probability /of failure/. I believe that to reconcile the two schools is to insist that a good design of the downstream DS zone of a dam must incorporate a pretested condition and consequent FG. In short, if the constructionperiod pore pressures affecting DS sliding instability are planned and made to be somewhat worse than those that will be introduced by the reservoir filling, and if soil shear strength only improves with time, the FG = 1.4 /say/ does indeed meet the deterministic design requirement.

ABANDONING CALCULATION OF SLOPE STABILITY, AND ADOPTING CALCULATIONS OF INCREMENTAL ACTIONS - EFFECTS

Stress-strain-time trajectories have, over the past 40 years, been shown to be important in conditioning behaviour parameters in laboratory tests. Of course, there has been a concomitant effort /successful/ to prove that the limit-strength condition, the effective stress Mohr envelope, does not change. Although some significant part of this demonstration of negligible hysteresis effects may be due to the modern rush-testing /automated, etc./ and consequent elimination of secondary compression and long-term effects, the "constant Mohr envelope" has established itself as an entrenched belief. And this has led to an unfortunate implication, that stability analyses can be profitably carried out as before. Since conventional stability analyses, assuming single-step "gravity-on" application of full deviator stress starting from isotropic conditions, and relinquishing long-term benefits, both accumulate on the conservative side, the design computation has rightfully proven acceptable, except in unusual soils /e.g. Scandinavian quick clays, etc./.

What must be emphasized is that the Mohr strength equation in effective stresses may remain essentially unchanged, but, because of significantly different pore pressures generated, the position at which the different stress-trajectory tests meet the envelope are quite different. Therefore stress-strain brittleness or ductility, and pore pressure estimates, should be highly dependent on trajectories: and, therefore, there cannot be, by conventional parameters and calculations, as satisfactory a prediction of slope stability behaviour as claimed.

We could draw the parallel with the wide differences detected, right from the start of finite element analyses /Clough 1966/, between gravity-on embankments and the multiple-step constructed embankment. And we could merely reason, quite generally, that if you start from the coordinate axes /O, O/ and apply a total change to a point /x, y/, your errors and dispersions must be much greater than if you approach the final /x, y/ condition incrementally.

In short, there should not be any stability analyses as such /of the rigid-block gravity-on conventional type/, but analyses of incremental stabilizing or unstabilizing causes and effects applied one at a time. This is the natural consequence, transferred to the prototypes, of the stress-strain-time trajectory concepts of laboratory testing of Taylor /1947/, much developed by Lambe. For the designer it gives a much better feel for relative importance of different parameters and causes. At any rate, for the development of behaviour histograms of Statisfaction Indices of slopes, for correlation with nominal FS values, it will be much more fruitful to analyse Δ/Satisfaction Indices/ vs. Δ FS values.

In professional practice it is more than 15 years since I have stopped conducting so-called stability analyses, and substituted them by analyses of stabilizing and unstabilizing effects.

BAYESIAN ADJUSTMENTS APPLIED TO INCREMENTAL ANALYSES

In all analyses /either for design computations, or for behaviour-prediction calculations/ it is very profitable to associate incremental cause-effect calculations with Bayesian adjustments of posterior vs. prior best-estimate probabilities. I have denominated this concept the quantification of Peck's "Observational Method" /Rankine Lecture 1977/, because I believe that man's acquisition of knowledge and culture is, subconsciously, by a Bayesian process of natural selection of decisions/actions/ effects.

The case of soft-ground tunneling lends itself very well to the discussion because in tunneling /design and construction/ there is first a very major effect /execution effect/ associated with the start /complex combination of equipment + methods + workmanship + subsoil etc./ and subsequent adjustments occur along the line, as conditions change or treatments are applied.

So very dominant are the effects of equipment and procedures, that it is quite spurious to lump together into a single presumed statistical universe the case-histories of soft-ground tunnels in London, Stuttgart, Tokyo, Lisbon, etc., or even of London 1950, London 1962, London 1974, or of Paris 1972 with chemical grouting, or with compressed air, or with deep-well groundwater drawdown, or with bentonite shield, and so on.

The first step for improvement of our knowledge comes from a deterministic separation of each tunneling project as a distinct physical fact, a distinct statistical universe. Then, with this most dominant starting fact maintained constant, what we want to know, and what the designer and contractor most want to know to orient their progress along the line, is how to adjust the varying conditions. Thus, for instance, as the tunnel advances in a sand from a condition 10 m below groundwater to 15 m below groundwater, what should be the predictable change: or what should be the predictable change if, all factors maintained constant, at a constant 10 m below groundwater, the tunnel advances from a loose sand condition to a dense sand condition; or, in the loose sand condition, if the compressed air is changed from 0.8 to 1.2 atmospheres, and so on.

Since we recognize that infinite are the factors affecting the problem, as for any problem $x = f\ /a, b, c\ldots\infty/$, the fruitful principle of laboratory testing was and is to work with partial differentials, $\partial x/\partial a$ with all other parameters maintained constant, then $\partial x/\partial b$ with all others constant, and so on. How can we expect success in the field and prototype laboratory, unless we reasonably respect the same rule of investigating /collecting statistical data on/ partial differentials? To the contractor, who must decide what incremental treatment he must apply in facing new troubles as he advances, the only need is of information on probable incremental effects. For these he must start with a statistical prior probability based on past experience; and, as rapidly as possible, after the first advances, he must adjust his educated-guess prior probability to a new best-guess posterior probability /Bayesian statistics, Observational Method/.

The important point to emphasize is that we must shun statistics at random, and choose to apply statistical adjustments to our reasonable theories. Each effect of groundwater, stresses and strains, instability, etc., can be analysed in a preliminary nominal degree through our theories: these theoretical treatments must be applied as the background for our search of the most probable statistical regressions. Of course, the temporary application of a presumed theory does not preclude concluding that it is not satisfactory, and consequently revising it, or even proposing an entirely different one: what cannot be condoned is the attempt to extract conclusions from data at random and spurious statistics, without any theory, however nominal, or any design and purpose, since such efforts prove sterile, and might even lead to dangerous conclusions /Revista Geotechnica, Lisbon 1983, ITA, Brussels 1983/.

In short, we must postulate some theorizable hypothesis, collect controlled data of partial differentials, analyze it statistically, adjust it gradually by Bayesian decision theory. That, in summary, is, to my mind, the most profitable avenue for progress in geotechnical design and construction.