Structural Eng. / Earthquake Eng. Vol. 6. No. 1, 201s- 205s. April 1989 / Discussion-Closure
Japan Society of Civil Engineers (Proc. of JSCE No. 404/I-11)

UPDATING FIRST-AND SECOND-ORDER RELIABILITY ESTIMATES BY IMPORTANCE SAMPLING

(*By Munehisa FUJITA and Rüdiger RACKWITZ*)

Structural Eng./Earthquake Eng. Vol. 5, No. 1, 31s~37s, April 1988
(Proc. of JSCE No. 392/I-9)

Discussion——————————*By Ulrich BOURGUND* (*JSPS Research Fellow, Kyoto University*)

Congratulations to the authors for the clarifying discussion on different importance sampling procedures for possible application in reliability analysis. Despite of the well documented comparative calculations for the presented procedures, some remarks should be made in order to provide further criteria for the selection of the most suitable method in general reliability estimation.

(1) The discussion about efficiency and accuracy of the presented importance sampling procedures is based on the assumption, that the design point can be calculated precisely. But obviously there is no guarantee, that any procedure to determine the design point calculates the right or exact one. Therefore this important condition has to be emphasized especially with respect to the presented method B and C, which might be rather sensitive to wrong or unidentified design points. Exclusively the so called direct method—requiring the coordinates of the design point—seems to be rather insensitive to wrong design points and will yield in this case an increased statistical error only. Furthermore this method enables calculations in the original space [16),17)], which gives the opportunity to use available information about the joint destribution.

At this point it might be useful to mention, that there are importance sampling methods, which do not require any design point calculation, since they are based on adaptive algorithms (e. g. 18), 19)) and therefore appear to be rather powerful in general reliability analysis. These methods could be applied efficiently even in the presence of several limit state functions (see 20)).

The following example, reflecting the investigation of the yield moment of a cantilever beam due to a single load, is calculated in order to emphasize the forementioned general remarks. The limit state function of this somehow realistic problem is given by

$$g(x)=x_1x_2-PL$$

where the basic variables x_1, x_2 might be considered as yield stress and modulus of plasticity, respectively. The loading P ($P=14.614$ kN) is considered as a deterministic dead load applied to the cantilever beam of length $L=10.0$m. All basic variables are assumed to be normal distributed, with mean values and standard deviations given as $\overline{x}_1=78\,064.42$ kNm, $\sigma_1=11\,709.663$ kNm, $\overline{x}_2=0.0104\text{m}^3$, $\sigma_2=1.56\,E-3\ \text{m}^3$, respectively, which is considered as a reasonable choice in case of a reinforced concrete structure. Calculating the failure probability of the limit state function two problems will show up. First two design points can be detected, and second the calculated design point is sensitive to start values of most calculation procedures. Furthermore if the start values are the mean values—which actually is the usual choice—the correct design point is not found by most of the known calculation procedures. Usually in this case the design "point" is calculate as the largest distance. It would be interesting to know how in this situation (several design points and/or difficultes to calculate the design point) the discussed method B and C could be applied . Method A[17] for this example yields $P_f=0.133E\text{-}6$, COV[%]=23.0(512 Simulations) while Adaptive Sampling[18]—requiring no design point calculation—gives $P_f=0.1441E\text{-}6$, COV[%]=18.9(514 Simulations). The exact result was calculated by conditional Integration[19]—which is very efficient for problems up to 20 variables—with $P_f=0.145137E\text{-}6$, COV[%]=0.01(5287 function calls). It should be noted, that conditional integration already yields a rather precise result with about 1 000 function calls $P_f=0.14506E\text{-}6$, COV[%]=1.55. It should be mentioned, that of course the characteristic of the presented explanatory example is not a general one, but at least reveals some properties which have to be expected in general reliability analysis and are often unknown in advance.

(2) Since the variance of the estimated failure probability does not only depend on the shape of the limit state surface but also on the shape of the simulation density in the failure domain[21),22)] the original distribution of the basic variables might be used as simulation density, reducing the variance of the result considerably. In this way the original problem can be treated efficiently avoiding the somehow "artificial" nonlinearity introduced by transforming the limit state surface into the standard normal space. From this point of view the selected example might be not considered as an extreme one, since the nonlinearity is introduced by the transformation process only, while the problem itself or the physical background is linear.

(3) For the discussion of the presented sampling methods a rather low number of simulations for the presented example has been selected. In this situation it should be noted, that the quality of the random number generator has some influence on the quality of the result, especially in cases of increased dimension[23),24)]. Therefore the variance of the results has to be expected rather fluctuating, since the quality of the random number generator especially will effect the tails.

It might be important to note, that all failure probabilities calculated by method B for the upper tail (see Table 1) are smaller than the exact one. According to 25), this might be an indicator of an underbiased or overbiased situation. It would be interesting to know, if the authors have performed any checks on this phenomenon. For the calculation of the upper tail in Table 3 again the failure probability of method B is smaller than the exact one. Since a common feature of an underbiased or overbiased situation is at the same time a small estimate of variances the comparison between the thre procedures has to be performed with caution.

(4) With respect to the SORM approach as basis for method C it should be noted, that in the presence of some noise due to involved calculations for the limit state function the application of the forward difference scheme might yield an unacceptable large error. In this situation the central difference scheme has to be applied which considerably increases the numerical effort to (see eq. (15)):

SORM $K_1(n+1)+2(n-1)^2$

Method C $K_1(n+1)+2(n-1)^2+2NK_2$

(5) Furthermore due to the writers experience the factor K_1 for the assesment of the numerical effort to find the design point (see eq. (15)) cannot be typically considered about 10. This value depends very much on the numerical procedure selected as well as on the required accuracy and in general varies between 5 and 100[26)]. Some procedures even have to be considered as rather unstable since they do not find the design point in the presence of noise[27)] introduced by involved calculations for the evaluation of the limit state function. Unfortunately the authors do not mention which kind of procedure has been selected. To the writers experience and in accordance with investigations reported in the literature[26),27)] a sequential quadratic programming technique seems to be most powerful, yielding $5(n+1)$ function calls for the presented examples.

(6) In order to improve the objectivity of the comparison, to the writers opinion the necessary CPU time should be listed in Table 3, since the number of function calls is not the only indicator for the assessment of efforts, because of big differences in numerical procedure of the investigated 3 methods. Furthermore it should be emphasized, that the failure probability and the estimated COV are conditional

estimates for method B and C, assuming a correct calculation of the design point (see section (1)). From this point of view the conclusion that there is a preference to method B over method A may be not quite obvious, since there is no reliable check if the right or all design points are calculated.

(7) With respect to larger structural systems even FORM/SORM Methods might require prohibitively large computational efforts since one evaluation of the limit state function e. g. requires a complete finite element analysis. In this situation the only practical tool for reliability estimation might be a combination of importance sampling and response surface approach, which could be efficiently demonstrated in 28).

ACKNOWLEDGMENT : The writer wishes to acknowledge the support of the Japanese Society for the Promotion of Science (JSPS) which enables this investigation and the scientific discussion with Japanese researchers at Kyoto University. Furthermore special thanks to Professor H. Kameda host scientist of the writer at Urban Earthquake Hazard Research Center, Disaster Prevention Research Institute.

REFERENCES

16) Schuëller, G. I. and Stix, R. : A Critical Appraisal of Methods to Determine Failure Probabilities, J. of Structural Safety, Vol. 4/4, pp. 293-309, 1987.
17) Bourgund, U. : Bucher, C. G. : Importance Sampling Procedures Using Design Points (ISPUD)—A User's Manual, Report No. 8/86, Institute of Engineering Mechanics, University of Innsbruck, Austria, Oct. 1986.
18) Bucher, C. G. : Adaptive Sampling—An Iterative Fast Monte Carlo Procedure, in Report No. 12/87, Institute of Engineering Mechanics, University of Innsbruck, 1987.
19) Ouypornprassert, W. : Conditional Integration : An Efficient Adaptive Numerical Integration for Reliability Analysis, Institute of Engineering Mechanics, University of Innsbruck, Internal Working Report No. 25, Dec. 1987.
20) Chen, T., Schuëller, G. I. and Bourgund, U. : Reliability of Large Structural Systems Under Time Varying Loads, Proc. of Special ASCE Conference, Blacksburg, Virginia, U. S. A., May 1988.
21) Rubinstein, R. Y. : Simulation and Monte Conte Carlo Method, John Wiley & Sons, New York 1981.
22) Bourgund, U., Ouypornprassert, W. and Prenninger, P. H. P : Advanced Simulation Methods for the Estimation of System Reliability, Internal Working Report No. 19, Institute of Engineering Mechanics, Univ. Innsbruck, 1986.
23) Bourgund, U. : Nichtlineare zuverlaessigkeitsorientierte Optimierung von Tragwerken unter stochastischer Beanspruchung, Ph. D. Thesis, Institute of Engineering Mechanics, University of Innsbruck, Dec. 1987 (in German).
24) Knuth, D. E. : The Art of Computer Programming : Seminumerical Algorithms, Vol. 2, Addision-Wesley, Reading Massachusetts, 1969.
25) McGrath, E. J. and Irvıng, D. C. : Techniques for Efficient Monte Carlo Simulations, Vol. III, Variance Reduction.
26) Schuëller, G. I. : Bucher, C. G. : Bourgund, U. and Ouypornprassert, W. : On Efficient Computational Schemes to Calculate Failure Probabilites, Proc. of US-Austrian Joint Seminar on "Stochastic Structural Mechanics", Y. K. Lin, G. I. Schueller (Eds.), 1987, Boca Ration, U. S. A., pp. 388-411, Submitted to Journal of Probabilistic Engineering Mechanics.
27) Liu, P. L. and Der Kiureghian, A. : Optimization Algortihms for Structural Reliability Analysis, Rep. No. UCB/SESM-86/09, University California, Berkley, July 1986.
28) Bucher, C. G. and Bourgund, U. : Efficient Use of Response Surface Methods, Institute of Engineering Mechanics, University of Innsbruck, Report No. 9/87.

(Received August 1 1988)

Closure——————————*By Munehisa FUJITA and Rüdiger RACKWITZ*

The authors appreciate very much the opportunity to clarify further the subject due to the interesting remarks of the discusser. The main and repeated criticism of the discusser is that the methods compared in our paper set out from FORM/SORM concepts. FORM/SORM concepts imply that a probability distribution transformation into a normal space is necessary and that the design point exists and can be determined uniquely. As an alternative the discusser proposes a so-called adaptive scheme to locate the important region. This alternative is claimed to work also in the original space of basic variables, to be able to locate the important region and to be powerful. These claims are illustrated by a numerical example. It is not possible here to check the validity of these claims in general. At least, the very illuminating example chosen by the discsser is absolutely inadequate for this purpose because his alternative just fails. It is , in fact, an example for the inadequacy of all importance sampling schemes known to us so far. The parameters are chosen such that the failure curve in the standard normal space is very colse to a circle in the third quadrant. More precisely, it is a hyperbola but only one branch of it is physically meaningful. On this curve a maximum point with $\beta=5.428$, identified as such by SORM as it has a radius of curvature of $R=4.00<\beta=5.428$ pointing to the origin and two symmetrical minimum points with $\beta=5.333$ (and radius of curvature $R=9.48>\beta=5.333$ also pointing to the origin) can be found. Had the discusser in his method—A—run not stopped at only 512 simulated points, he would have found that the results are not converging, i. e. the COV jumps up from time to time while it decreases slowly in between the jump events. A jump occurs if a simulated point falls far off the rest and the quotient of the two densities (see eq. (5)) reaches values of several orders of magnitude. Analysis of the example with the methods B and C show a similar behavior. Any other method of importance sampling must show this behavior unless the spread of the sampling density is chosen so wide that the efficiency of the method is almost down to crude Monte Carlo. In this example there simply is no explicit important region except that one can focus on the third quadrant. This is the reason why the existence of a unique design point is assumed in our paper and with it a unique important region. Also, the transformation into the standard space is mandatory because only the corresponding standard normal density has the property of rotational symmetry guaranteeing that a design point is (locally) unique. The use of the curvature properties in this point enables to verify that it is a minimum point. Besides, only the properties just mentioned allow to prove the mathematically strong asymptotic SORM—results. Finally, it is a property of most search algorithms for the design point that they converge, if they converge, significantly slower than usual in cases as given by the discusser. This slower convergence rate can be used as an indication of possible problems. Thus FORM/SORM gives warnings if there are problems and importance sampling is only used to update usually already accurate results.

If there are several minimum points with almost equal geometric safety index one is in serious trouble. Sometimes one can cure the problem in formulating the physical context in terms of "failure modes" such that each one of those has a unique design point. Then, each "failure mode" is treated separately and the overall failure probability is the probability of the union of the "failure modes". Otherwise any importance sampling method is prone to failure including the one proposed by the discusser. It is even dangerous to use it because it may show convergence by a decreasing COV for not too large sample sizes (see the discussion) despite a possibly strong bias in the probability estimate but without giving any warning that this is so. That warning can only be expected for relatively large sample sizes. For example, in our reanalysis of the discusser' s example with method A the jump in the COV occurred at the 8981-th sampling point. The number when this occurs, of course, depends on the particular random number generator. In our opinion the only way out of such situations is numerical integration, crude Monte Carlo or so-called spherical sampling.

Fortunately, such degenerate cases are very rarely met in real world applications. The interested reader can convince himself easily by applying the Japanese standard[29] for the determination of the parameters in the discusser's example. With realistic parameters FORM and/or SORM and especially method C and the discusser's method yield very accurate and inexpensive results.

In non-degenerate cases the situation with pure importance sampling is even worse if for some reason a wrong region is chosen as the important region. Those methods then show fast convergence without giving any warning that the estimate has a possibly strong bias. As a matter of fact, pure importance sampling schemes have hidden fallacies in applications a few of which have been pointed out in the above. The conclusion, therefore, is that the important region must exist and must be identified as such for any importance sampling scheme and this is efficiently done by FORM/SORM which also provides clear indicators of possible problems. It needs to be demonstrated that the adaptive scheme mentioned by the discusser also has such capabilities. If there is no clear important region totally different methods are in order a few of the possibilites are mentioned before.

With respect to the rest of the discusser's remarks we only mention that a well-tested random number generator was used[30] and that it is not obvious how to rank one search algorithm above others in general. Many of them have special features such as fast or sure but slow convergence or robustness against numerical noise. But none combines all of these features. In our own FORM/SORM programs the type of algorithm can be chosen according to the special needs in an application. But the last subject was on purpose not a topic of our paper.

REFERENCES

29) Standard Specification for Design and Construction of Concrete Structures, Part 1 (Design), JSCE, 1986.
30) Dalgas-Christiansen. H : Random number generators in several dimensions, Rep., Institute for Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, 1975.

(Received February 25 1989)