



New importance sampling estimation schemes for sensitivity analysis and rare event estimation

Nouvelles méthodes d'estimation par échantillonnage préférentiel pour l'analyse de sensibilité et l'estimation d'évènements rares

Soutenance de thèse – Julien DEMANGE-CHRYST

Lundi 18 novembre 2024 à 14h

Amphithéâtre SCHWARTZ, Université Paul Sabatier

Institut de Mathématiques de Toulouse, 118 route de Narbonne 31400 Toulouse

Devant le jury composé de :

Sébastien DA VEIGA	ENSAI	Rapporteur
Arnaud GUYADER	Sorbonne Université	Rapporteur
Béatrice LAURENT-BONNEAU	INSA Toulouse	Examinatrice
Nicolas BOUSQUET	EDF Lab Saclay	Examineur
Erwan SCORNET	Sorbonne Université	Examineur
François BACHOC	Université Toulouse III – Paul Sabatier	Directeur de thèse
Jérôme MORIO	ONERA Toulouse	Co-directeur

Résumé

Dans de nombreux domaines scientifiques, l'évaluation de la fiabilité et plus généralement l'étude d'un système complexe sont réalisées grâce à un modèle numérique visant à simuler le plus fidèlement possible son comportement. L'intérêt d'une telle modélisation est d'être en mesure d'étudier le système de manière numérique et non expérimentale. Cette approche est particulièrement bénéfique dans le cas d'un système critique où toute défaillance peut avoir d'importantes conséquences. Mathématiquement parlant, un tel modèle est généralement considéré comme une fonction déterministe de type boîte noire et est très coûteux en temps à évaluer. Puis, le système complexe est soumis à des incertitudes aléatoires en entrées qui sont modélisées dans un cadre probabiliste par une distribution multidimensionnelle. La prise en compte de ces incertitudes en entrées ainsi que leur propagation au travers du modèle forment la méthodologie de la quantification d'incertitudes. Dans ce contexte, nous allons dans cette thèse nous intéresser à trois questions majeurs. Tout d'abord, l'analyse de sensibilité, assimilée à une propagation d'incertitudes inverse, consiste à déterminer comment l'incertitude en sortie du modèle peut être attribuée aux différentes sources d'incertitudes en entrée. Puis, l'analyse de fiabilité consiste à estimer la probabilité de la défaillance du système, qui est un évènement rare défini par des considérations physiques. Enfin, l'analyse de sensibilité fiabiliste combine les deux études précédentes et fournit une compréhension plus approfondie de la défaillance. Chacune de ces trois analyses requiert le calcul de mesures quantitatives adaptées permettant d'apporter une réponse à l'objectif de l'étude, et leur estimation précise renferme alors un enjeu majeur afin de correctement évaluer les risques du système. Des méthodes d'estimation spécifiques ont été développées dans la littérature, mais ne répondent pas à toutes les embûches qui peuvent être rencontrées en pratique. C'est pourquoi nous introduisons dans cette thèse de nouvelles méthodes par échantillonnage préférentiel visant à réduire l'erreur d'estimation des indices de sensibilité et de la probabilité de défaillance dans le cas où la défaillance est un évènement rare et/ou en grande dimension. Dans un premier temps, la procédure d'échantillonnage commune pour l'estimation des indices de Sobol pour l'analyse de sensibilité n'est pas optimale dans l'optique de réduire au maximum l'erreur de leur estimation. Notre première contribution est alors la construction de nouveaux estimateurs visant à minimiser un critère d'optimalité de manière adaptative en couplant les méthodes de réduction de variance de l'échantillonnage

préférentiel et des variables de contrôle. Dans un second temps, les méthodes d'estimation de la probabilité de défaillance spécifiques à la grande dimension ne sont généralement pas très flexibles dans le sens où de fortes connaissances préalables sur le problème et la défaillance sont requises pour qu'elles soient pertinentes et permettent d'obtenir une faible erreur. Pour y remédier, nous proposons de tirer avantage de la flexibilité et de la robustesse face à la dimension d'un auto-encodeur variationnel, un modèle génératif basé sur des réseaux de neurones, en l'adaptant et en l'utilisant dans un contexte d'échantillonnage préférentiel. Dans un dernier temps, la méthode d'estimation existante des indices de Shapley pour l'analyse de sensibilité fiabiliste ne prend pas en compte les contraintes liées au fait que la probabilité de défaillance soit potentiellement très faible. Notre dernière contribution est alors de réécrire ces indices par échantillonnage préférentiel et nous montrons que les nouveaux estimateurs correspondant permettent de considérablement réduire l'erreur d'estimation lorsque la densité auxiliaire de tirage est choisie judicieusement. Pour finir, nous illustrons l'intérêt pratique de ces développements sur des cas tests académiques.

Mots clés

Quantification d'incertitudes, Analyse de sensibilité, Estimation d'évènements rares, Échantillonnage préférentiel, Auto-encodeur variationnel, Indices de Shapley.

Contact pour obtenir le lien de connexion à distance : jerome.morio@onera.fr