

FAUT-IL REJETER LE BOOTSTRAP ?

André Cabarbaye
CNES / CAB INNOVATION
3 rue de la Coquille – 31500 Toulouse
Tél. 05 61 54 68 08
andre.cabarbaye@cnes.fr /
andre.cabarbaye@cabinnovation.fr

Julien Faure
Centre National d'Etudes Spatiales (CNES)
18 av. Edouard Belin – 31401
Tél. 05 61 27 47 19
julien.faure@cnes.fr

INTRODUCTION

Les codes de calcul interviennent de plus en plus dans le dimensionnement des systèmes. Aussi, la maîtrise de leurs incertitudes constitue une préoccupation majeure des concepteurs ainsi que des organismes de sûreté et de certification. Ces incertitudes ont des origines multiples : méconnaissance des phénomènes physiques, limitation et représentativité des données expérimentales, dispersion et corrélation éventuelles des données d'entrée, faiblesse des modèles numériques, etc.

Ainsi dans le domaine spatial, les satellites sont traditionnellement dimensionnés pour assurer une mission de référence dans les pires cas de fonctionnement (en fin de vie, durant les éclipses, etc.). Mais cette approche, qui simplifie la validation des systèmes en limitant les vérifications à un cas de référence unique, conduit à un surdimensionnement général et à l'impossibilité de bénéficier de certains effets de seuil tel que l'emploi de lanceurs de plus faible capacité. Aussi envisage-t-on des dimensionnements probabilistes qui seraient valider à partir de résultats de simulation en considérant les aléas de toute nature portant sur la mission, les conditions orbitales et l'état du système à l'instant courant (défaillances et dégradations à bord).

Mais comment définir le nombre de simulations nécessaires à la démonstration de la tenue des objectifs ? C'est pour répondre à cette question que nous nous sommes intéressés à la méthode du Bootstrap, qui est actuellement préconisée par un groupe méthodologique de l'OCDE (Programme BEMUSE [1]) pour démontrer qu'en cas de perte de réfrigérant du circuit primaire d'un réacteur nucléaire, la température de gaine des crayons combustibles ne dépassera pas une température critique dans 95 % des cas

Proposée à la fin des années 70 par Bradley Efron [2], le Bootstrap est une méthode de sur échantillonnage sensée améliorer les estimations statistiques. A partir d'un échantillon original, correspondant par exemple aux résultats d'un certain nombre de simulations d'une mission spatiale, la méthode du Bootstrap consiste à générer un grand nombre d'échantillons fictifs, par tirage aléatoire avec remise, sans recourir à de nouvelles observations. L'objectif est ainsi de mieux exploiter l'information contenue dans l'échantillon original même si aucune information nouvelle n'est ainsi créée. Mais si ce sur échantillonnage a, peut-être, le pouvoir d'améliorer certains estimateurs par rapport à ceux obtenus directement à partir de l'échantillon initial, nous doutions qu'il puisse améliorer significativement la précision d'un quantile à 95% et surtout autoriser une réduction de la taille de l'échantillon initial pour un même objectif à atteindre. Aussi avons-nous cherché à tester la capacité réelle de cette méthode.

PROTOCOLE DE TEST

La démarche de validation mise en oeuvre pour tester cette méthode est simple et pragmatique. Nous avons considéré un phénomène physique régi par une loi théorique donnée dont on connaît a priori la valeur des quantiles ainsi que d'autres paramètres caractéristique (moments, etc.). Nous avons alors tenté d'estimer ces mêmes paramètres par la méthode du Bootstrap à partir d'un grand nombre d'échantillons générés par simulation de Monte-Carlo de la loi théorique.

Pour cela, nous avons procédé au tirage aléatoire d'un échantillon de 59 valeurs distribuées suivant la loi théorique à partir duquel nous avons généré 250 échantillons « bootstrapés » pour estimer le quantile à 95 % et les autres paramètres caractéristiques de la distribution, que nous avons comparés à leur valeur théorique. Répétant 2000 fois l'opération, nous avons obtenu les fonctions de répartition de l'écart entre les valeurs estimées et théoriques. De simples lois normale et lognormale ont été choisies comme loi théorique.

CONCLUSION

La confiance réelle de l'estimation du quantile à 95 %, réalisée par la méthode du Bootstrap, n'a pas dépassée 50 % au cours de cette évaluation, bien loin des valeurs escomptés (95% pour les participants du programme BEMUSE).

Aussi, nous considérons que la méthode du Bootstrap doit être rejetée pour estimer un quantile même si son intervalle de confiance peut être très resserré puisque le nombre d'échantillons générés par tirage aléatoire peut être augmenté à loisir.

L'apport du Bootstrap s'est également révélé décevant pour évaluer d'autres estimateurs tels que la moyenne ou l'écart type. En effet, les estimateurs obtenus par cette méthode s'avèrent quelque peu fictifs car ils ne correspondent pas à ceux de la population mère mais à ceux d'une population virtuelle pouvant être générée à partir d'un échantillon.

Cependant, nous avons vérifié, au cours de ces mêmes tests que la méthode de Wilks [3], datant de plus d'un demi-siècle et traditionnellement utilisée en thermo-hydraulique nucléaire, donne d'excellents résultats pour l'estimation d'un quantile. Elle sera utilisée prochainement dans le domaine spatial pour valider des dimensionnements de nature probabiliste à partir de résultats de simulation stochastique. Remplaçant des évaluations déterministes en pire cas, ce type de dimensionnement nous apparaît en effet susceptible de limiter les marges inutiles et de valoriser au mieux les moyens existants.

Références

- [1] JOUCLA, J., CHOJNACKI, E., Expertise des méthodes d'évaluation des incertitudes du code CATHARE dans le programme international OCDE BEMUSE, IRSN – Rapport scientifique et technique, 2006.
- [2] EFRON, B., TIBSHIRANI, R. J., An Introduction to the Bootstrap, Chapman and Hall, 1993.
- [3] WILKS, S. S., « Determination of Sample Sizes for Setting Tolerance Limits », The Annals of Mathematical Statistics, Vol.12, pp. 91-96, 1941.