



La Méthode de Wilks :

Utilisation incorrecte pour les études de sûreté

par Bernard Beauzamy

Janvier 2016

Abstract

La "méthode de Wilks" consiste en un énoncé probabiliste extrêmement simple (niveau Terminale) : étant donné un phénomène aléatoire quelconque, si on se fixe des boîtes d'égale probabilité, on finira certainement par tomber dans chacune d'entre elles, au bout d'un nombre suffisant de "runs" ou d'expérimentations. Cette méthode est souvent utilisée dans les études probabilistes de sûreté, parce qu'elle permet de déterminer le nombre minimum de runs nécessaires (pour une expérience, pour un code de calcul), afin d'évaluer un risque avec un niveau de sécurité jugé suffisant.

Mais l'utilisation qui en est faite est fondamentalement incorrecte : les ingénieurs l'appliquent aux différentes lois qu'ils ont arbitrairement introduites sur les paramètres d'entrée du code, alors qu'il faudrait l'appliquer à la loi de sortie du code. L'utilisation de cette méthode pour les études probabilistes de sûreté est absolument à proscrire.

Utiliser, de manière aveugle, quelques milliers de runs d'un code de calcul qui dépend de dizaines de paramètres complexes, en espérant que Dieu, dans sa bienveillance, aura la gentillesse de les guider vers les situations à risque, permettant ainsi de les identifier, relève de la naïveté, de l'aveuglement, de la paresse. Un aveugle ivre, jetant au hasard une fléchette en espérant qu'elle franchira une porte à peine entrebâillée, aurait des milliards de fois plus de chances de succès.

Nous donnons ci-dessous un exemple très clair et complètement indiscutable, qui montre pourquoi la méthode de Wilks est utilisée de manière fautive. Mais, au-delà de cette méthode, c'est l'utilisation des probabilités en général, pour les approches de sûreté, qui est à reprendre : il est fondamentalement absurde de s'en remettre au hasard pour l'investigation d'un phénomène que nous ne comprenons pas.

I. Présentation générale de la méthode

La méthode est présentée dans de nombreux articles. Retenons l'un d'eux : "Estimation d'un quantile concourant à la maîtrise d'un dimensionnement", présenté lors du 16ème Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement - Avignon 6-10 octobre 2008, communication 1Z-2 pages 1-6, par André Cabarbaye (CNES/CAB Innovation) et Roland Laulheret (CNES)

http://cabinnovation.drupalgardens.com/sites/cabinnovation.drupalgardens.com/files/fichiers/LBD16_1.pdf

La méthode de Wilks [1] est la méthode la plus employée à ce jour en thermo-hydraulique nucléaire pour estimer un quantile. Elle permet de déterminer le nombre minimum N de simulations nécessaires à l'obtention d'un majorant de la valeur d'un quantile $T_{\alpha,\beta}$.

Ce nombre est donné par la formule de Wilks, qui résulte directement de l'expression de l'intervalle de confiance d'une loi binomiale :

$$1 - \sum_{i=N-r+1}^N \binom{N}{i} p^i (1-p)^{N-i} \geq \beta$$

p est la probabilité qu'une simulation quelconque aboutisse à un résultat satisfaisant (inférieur à la valeur critique), et r est le rang des pires cas obtenus durant toutes les simulations. L'estimation du quantile $T_{95,95}$ est donnée par la valeur pire cas obtenue au cours de N simulations ($r = 1$) avec N donné par l'expression :

$$1 - \alpha^N \geq \beta, \text{ soit } \alpha^N \leq 1 - \beta, \text{ ou } N \geq \frac{\text{Ln}(\beta - 1)}{\text{Ln}(\alpha)}$$

$N \geq \text{Ln}(0,05) / \text{Ln}(0,95) = 58.4039748$. Soit $N \geq 59$.

Ce même quantile $T_{95,95}$ peut être estimé par la pire des valeurs obtenue à l'exclusion du pire cas ($r = 2$) avec N donné par l'expression :

$$1 - (N\alpha^N (1 - \alpha) + \alpha^N) \geq \beta, \text{ soit } N \geq 93,$$

ou par la pire des valeurs obtenue à l'exclusion des 2 valeurs pires cas ($r = 3$) soit $N \geq 124$.

II. Analyse mathématique de la formule

Nous nous intéressons ici à l'estimation du quantile $T_{95,95}$ donnée par la valeur pire cas obtenue au cours de N simulations ($r = 1$), où N est donné par l'expression $1 - p^N \geq \beta$. Voyons d'où vient cette formule : avant d'en critiquer l'usage, il faut bien comprendre le contenu.

La "Méthode de Wilks" consiste en un énoncé probabiliste extrêmement simple :

Soit X une variable aléatoire à valeurs réelles. Divisons l'ensemble des valeurs prises par X (peu importe en quoi elles consistent) en n sous-ensembles de même probabilité (en anglais "bins", souvent traduit par "boîtes"). Evidemment, la probabilité de chaque boîte est $p = \frac{1}{n}$. Désignons par B_1, \dots, B_n ces boîtes. Choisissons l'une d'entre elles, par exemple la dernière, B_n . Supposons que nous fassions N tirages indépendants de la variable X (peu importe quelle est la loi de cette variable). Alors la probabilité de ne jamais tomber dans B_n au cours de ces N tirages est :

$$q = (1 - p)^N$$

Ce résultat est évident : la probabilité de ne pas tomber dans B_n lors d'un tirage est $1 - p$ et donc la probabilité de ne jamais y tomber en N tirages est $(1 - p)^N$, puisque les tirages sont indépendants.

Comme $(1 - p)^N \rightarrow 0$ lorsque $N \rightarrow \infty$, nous sommes de plus en plus sûrs de toucher toutes les boîtes, lorsque le nombre de tests augmente. Fixons-nous un seuil β , traditionnellement 95% ; nous constatons que notre probabilité de toucher la dernière boîte en N runs dépasse ce seuil dès que :

$$(1 - p)^N \leq 1 - \beta$$

On retrouve la notation de l'article cité au paragraphe précédent en posant $\alpha = 1 - p$. Avec les valeurs numériques prises ici, à savoir $p = \frac{1}{20}$, on a $\alpha = 0.95$, et avec $\beta = 0.95$, on trouve effectivement $N \geq 59$.

III. Application de la méthode à l'analyse de sûreté

Celle-ci est généralement faite sous la forme suivante.

On dispose d'un code de calcul, supposé représenter un phénomène complexe. Nous prendrons l'exemple du code de thermohydraulique "CATHARE" (Code Avancé de THERmohydraulique pour les Accidents des Reacteurs à Eau, <http://www-cathare.cea.fr>, développé par Areva, le CEA, EDF et l'IRSN). Ce code décrit le comportement d'un réacteur en cas de grosse brèche dans le circuit réfrigérant primaire ; il retourne une température en fonction du temps et dépend d'un nombre très élevé de paramètres (40 ou davantage, selon les versions). C'est un code extrêmement complexe, qui a requis des dizaines d'années de développement. Le code est accepté par les Autorités de Sûreté, comme élément de validation.

Dans le livre Zeydina-Beauzamy "Probabilistic Information Transfer" [PIT], nous prenons l'exemple de ce code pour montrer comment "propager" l'information, de manière probabiliste, à partir de situations où les runs ont été faits à des situations où aucune information n'est connue.

Une autre situation pourrait être celle où l'on fait une vraie expérience physique, dépendant d'un grand nombre de paramètres. Pour des raisons de temps, de coût, de sécurité, on ne peut pas répéter l'expérience aussi souvent qu'on le voudrait.

La question fondamentale, pour les analyses de sûreté, est donc celle-ci : on dispose d'un petit nombre de runs (du code, de l'expérience), mettons 1 000 pour fixer les idées, et on voudrait en déduire une réponse à une question liée à la sûreté. Par exemple, pour CATHARE, un seuil est fixé arbitrairement à 1 200°C ; admettons que nos 1 000 runs aient tous donné des températures inférieures, voire bien inférieures, cela prouve quoi ? Que pouvons-nous remettre aux Autorités de Sûreté ?

La méthode de Wilks est utilisée, pour répondre à cette question, de la manière suivante :

Nous avons nos 40 paramètres. Sur chacun d'entre eux, nous choisissons arbitrairement une loi de probabilité. Par exemple, sur le premier, nous choisirons une loi uniforme sur tel intervalle ; pour le second nous choisirons une loi de Gauss de moyenne tant et de variance tant, et ainsi de suite jusqu'au 40^{ème}.

Par exemple, pour le paramètre X16 "pression accu", on retient une loi uniforme entre $m = 4\,137$ kbars et $M = 4\,385$ kbars. Ceci n'a rien de choquant : cela signifie que, les experts ne connaissant pas la valeur exacte que peut prendre le paramètre X16, lui attribuent n'importe quelle valeur entre ces bornes, avec égale probabilité. Cela reflète l'ignorance du phénomène précis.

Ces lois sont établies "à dire d'expert", c'est-à-dire que les experts considèrent que, en cas d'accident, le premier paramètre sera effectivement compris entre a et b , que le second sera plutôt concentré autour d'une certaine valeur, etc. Ces affirmations sont assurément discutables, mais elles ne sont pas absurdes et elles peuvent servir de base au raisonnement. Notre critique ne porte pas sur ce point.

Après quoi, on procède à des tirages aléatoires de chaque paramètre selon sa loi propre ; si pour le premier, on a choisi une loi uniforme, on fait un tirage selon une loi uniforme. Si pour le second on a choisi une loi de Gauss, on fait un tirage selon une loi de Gauss, et ainsi de suite pour les 40 paramètres. On répète l'ensemble de la procédure par exemple 1 000 fois (cela fait donc au total 40 000 tirages). On dispose ainsi de 1 000 "runs", chacun correspondant à une situation de fonctionnement du code (une situation de fonctionnement requiert 40 paramètres). Pour chaque run, le code fonctionne et retourne une température.

Les experts prétendent alors que, puisque on a fait plus de 59 runs choisis aléatoirement, la méthode de Wilks nous dit que nous avons plus de 95 chances sur 100 d'avoir détecté le quantile 95 des températures les plus élevées. Autrement dit, que nous sommes quasiment certains de la représentativité de nos essais : nos runs vont révéler la température la plus élevée.

Cette affirmation est entièrement fautive. L'erreur commise est celle-ci : il faudrait appliquer la méthode de Wilks à la réponse du code (qui n'est pas connue), et non aux lois d'entrée, qui sont fixées arbitrairement sur chaque paramètre.

IV. Où l'erreur se situe-t-elle ?

Nous allons bien expliquer ceci, sur des exemples de complexité croissante. Nous commençons par le cas où le code ne dépend que d'un seul paramètre. Il s'agit ici de revenir aux fondamentaux des probabilités. Ce n'est pas le code qui est en cause, ni la méthode de Wilks, mais la manière dont on l'applique, par ignorance des règles fondamentales des probabilités.

A. Cas d'un seul paramètre

Imaginons d'abord pour simplifier que le code dépende du seul paramètre X_{16} ; tous les autres sont oubliés, ou fixés à une valeur particulière.

On renormalise le paramètre X_{16} pour qu'il varie entre 0 et 1. Ceci est toujours possible, sans perte de généralité. On aura la température 1 000 pour $X = 0$ et la température 1 500°C pour $X = 1$.

Considérons d'abord un cas très simple :

– Cas 1 : la réponse du code est linéaire

Cela signifie que la température est fonction linéaire du paramètre X_{16} , et qu'on a une équation de la forme :

$$CT = 500x + 1000$$

Voyons en ce cas quelle est la loi de probabilité sur CT déduite de la loi uniforme sur x .

Pour tout λ , $1000 \leq \lambda \leq 1500$, on a :

$$\begin{aligned} P(CT \leq \lambda) &= P(500x + 1000 \leq \lambda) \\ &= P\left(x \leq \frac{\lambda - 1000}{500}\right) \\ &= \frac{\lambda - 1000}{500} \end{aligned}$$

puisque la loi sur x est uniforme. Par dérivation, on obtient la densité de probabilité de CT :

$f_{CT}(\lambda) = \frac{1}{500}$ si $1000 \leq \lambda \leq 1500$, 0 ailleurs : nous sommes donc en présence d'une loi uniforme. Nous constatons que si la réponse du code est linéaire, la loi uniforme sur le paramètre

se transforme en une loi uniforme sur la réponse, et on peut appliquer directement la méthode de Wilks. L'application faite est correcte en ce cas.

– **Cas 2 : la réponse du code n'est plus linéaire, mais polynomiale**

On prend cette fois une réponse de la forme :

$$CT = 500 x^5 + 1000$$

si bien que l'on a toujours les valeurs $CT = 1000$ pour $x = 0$ et $CT = 1500$ pour $x = 1$:

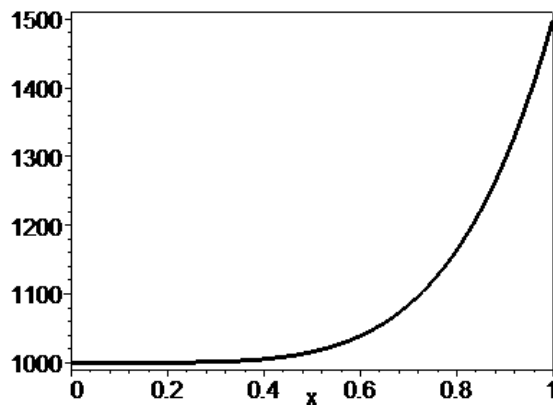


Figure 1 : réponse polynomiale

Voyons maintenant quelle est la loi de probabilité sur la réponse. Le calcul est le même que précédemment :

Pour tout λ , $1000 \leq \lambda \leq 1500$, on a :

$$\begin{aligned} P(CT \leq \lambda) &= P(500x^5 + 1000 \leq \lambda) \\ &= P\left(x \leq \left(\frac{\lambda - 1000}{500}\right)^{1/5}\right) \\ &= \left(\frac{\lambda - 1000}{500}\right)^{1/5} \end{aligned}$$

et par dérivation on obtient :

$$f_{CT}(\lambda) = \frac{1}{500} \left(\frac{\lambda - 1000}{500}\right)^{-4/5} \text{ si } 1000 \leq \lambda \leq 1500, 0 \text{ ailleurs ; voici le graphe de cette fonction :}$$

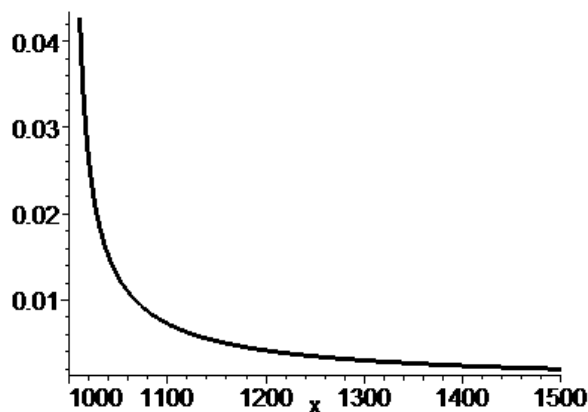


Figure 2 : la densité de probabilité associée

On constate qu'il ne s'agit plus du tout d'une loi uniforme : les faibles valeurs de température sont grandement privilégiées par rapport aux fortes valeurs. Autrement dit, si nous tirons au hasard des valeurs de X_{16} selon une loi uniforme, nous avons très peu de chances de tomber sur les fortes valeurs de la température. La méthode est complètement en défaut.

Nous avons tiré au hasard 59 valeurs du paramètre X_{16} , avec équiprobabilité, mais la réponse du code fait que nous n'abordons que très peu les valeurs élevées. La probabilité de détecter une température élevée sera donc très inférieure à ce que dit Wilks, lorsque la méthode est appliquée (à tort) sur la loi d'entrée.

B. Cas de la dimension élevée

1. Un exemple simple

Les choses ne font qu'empirer lorsque la dimension s'élève. Si on regarde d'abord le cas simple où le code ne dépend que de deux paramètres, mettons $CT = X_1 + X_2$, où X_1 et X_2 suivent toutes deux une loi uniforme et sont indépendantes, la réponse du code n'est pas une loi uniforme mais une loi triangle. La loi réponse se concentre de plus en plus lorsque la dimension augmente.

2. Un contre-exemple explicite

Voici maintenant un exemple explicite où les valeurs élevées de la température ne pourront jamais être détectées par la méthode de Wilks, telle qu'elle est utilisée actuellement. Ce contre-exemple, complètement évident, met bien en évidence l'erreur commise.

On suppose que la température, dépendant de 40 paramètres, est donnée par la formule :

$$CT = 2000 x_1 x_2 \cdots x_{40}$$

où chacun des x_i suit une loi uniforme sur l'intervalle $[0,1]$. La valeur maximale de la température est 2000°C , et elle obtenue si tous les x_i valent 1. On constate que le code est linéaire

par rapport à chaque paramètre pris séparément et qu'il est très régulier (aucune discontinuité d'aucune sorte).

Il est pratiquement impossible, par quelque simulation numérique que ce soit, de mettre en évidence une température supérieure à 1 000°C. En effet, $CT > 1\,000$ équivaut à

$$x_1 x_2 \cdots x_{40} > \frac{1}{2} \quad (1)$$

La propriété (1) implique que nécessairement, dans ce cas, tous les x_i doivent être $> \frac{1}{2}$. Or :

$$P\left(x_i > \frac{1}{2}, \forall i\right) = \frac{1}{2^{40}} \approx 0.9 \times 10^{-12} \quad (2)$$

Il faudra donc de l'ordre de 1000 milliards de runs pour rencontrer une situation où tous les x_i seront $> \frac{1}{2}$.

Ceci se comprend très bien intuitivement : si l'on tire 40 fois au hasard selon une loi uniforme entre 0 et 1, il est très vraisemblable que l'un des tirages au moins sera entre 0 et 1/2, et en ce cas le produit le sera aussi. La détection de la situation à risque est pratiquement impossible par une simulation numérique.

Remarque. – On peut calculer précisément la probabilité de l'événement (1). En effet, le produit de n variables aléatoires indépendantes, suivant une loi uniforme sur $[0,1]$, a pour densité (voir [2]) :

$$f(x) = \frac{1}{(n-1)!} \left(\text{Log} \left(\frac{1}{x} \right) \right)^{n-1}$$

V. En conclusion

Un code de calcul, dépendant de 40 paramètres, qui a nécessité des dizaines d'années de développement, est un objet fondamentalement complexe. Il est naïf et illusoire de croire que, en lançant 1 000 runs, ou un million de runs au hasard, Dieu aura la bonne volonté de guider nos fléchettes vers les endroits que nous souhaitons atteindre.

Si le code dépend de 40 paramètres et que chacun est discrétisé en 10 valeurs possibles, nous obtenons 10^{40} possibilités. L'exploration par 59 runs, ou par 1000, ou par un million, reste de toute façon infime, qu'on le veuille ou non, qu'on utilise Wilks ou non. Nous avons, avec nos fléchettes, des milliards de fois moins de chances d'atteindre notre cible que n'en aurait un aveugle ivre lançant une fléchette au hasard ; il espère qu'elle parviendra à franchir une porte à peine entrebâillée.

VI. Recommandations

Il est fondamentalement absurde et malsain de s'en remettre au hasard pour l'exploration d'un phénomène que nous ne comprenons pas. Bien au contraire, il est nécessaire de combiner l'expertise physique et l'exploration probabiliste locale, pour la valider. Concrètement, il faudra :

- Définir des scénarios, par exemple selon les conditions d'exploitation du réacteur ;
- Pour chaque scénario, définir des conditions accidentelles ;
- Pour chacune, par expertise, mettre en évidence les paramètres les plus influents ;
- Eliminer les paramètres les moins importants et se concentrer sur les paramètres les plus significatifs (voir notre livre [NMP]) ;
- Valider ce choix en procédant à des tirages aléatoires au voisinage des situations retenues ;
- Eliminer les paramètres qui sont sans influence, pour réduire la combinatoire du problème ; après avoir ainsi réduit la dimension, rechercher les configurations à risque, et concentrer les runs sur celles-là, pour les identifier le mieux possible ;
- Répéter ces raisonnements de manière à avoir la meilleure connaissance possible des paramètres significatifs et des situations critiques : concentrer les runs à la seconde génération sur les situations reconnues comme critiques à la première, et ainsi de suite.

VII. Références

[1] WILKS, S. S., Determination of Sample Sizes for Setting Tolerance Limits, *The Annals of Mathematical Statistics*, Vol.12, pp. 91-96, 1941.

[2] Carl P. Dettmann and Orestis Georgiou : Product of n independent Uniform Random Variables, *Statistics and Probability Letters*, 79 (2009) 2501–2503.

[NMP] Bernard Beuzamy : *Nouvelles Méthodes Probabilistes pour l'évaluation des risques*. Ouvrage édité et commercialisé par la Société de Calcul Mathématique SA. ISBN 978-2-9521458-4-8. ISSN 1767-1175, avril 2010.

[PIT] Olga Zeydina et Bernard Beuzamy : *Probabilistic Information Transfer*. Ouvrage édité et commercialisé par la Société de Calcul Mathématique SA. ISBN: 978-2-9521458-6-2, ISSN : 1767-1175, mai 2013.