Méthode de recherche des unités de mesure probablement utilisées dans un complexe mégalithique

Version 1.0 – avril 2017

Introduction

Situation initiale et problématique :
Au sein d'un complexe mégalithique, nous avons relevé les distances séparant tous les mégalithes les uns des autres. A partir de ces relevés, nous souhaiterions savoir si une ou plusieurs unités de mesures ont été utilisées pour disposer les mégalithes du complexe.

Plan de la méthode de résolution :
N'ayant aucun *a priori* sur la (les) unité(s) potentiellement utilisée(s), il va nous falloir comparer tous nos relevés à un grand nombre d'unités afin de voir si certaines distances sont des multiples entiers des unités testées. Pour cela, il va donc falloir, pour chaque unité de mesure testée :

- → calculer la probabilité théorique qu'une distance soit exprimable dans cette unité (p) ;
- → déterminer combien de distances relevées sont des multiples entiers de l'unité (k) ;
- → calculer la probabilité d'avoir au moins autant de coïncidences dans nos relevés ($P(X \geq k)$).

Enfin, il faudra classer les unités de mesures selon les $P(X \geq k)$ décroissants afin d'identifier celles qui sont beaucoup plus présentes que le hasard ne le permettrait.

But du document :
Ce document va présenter les méthodes pour déterminer les paramètres p , k et $P(X \geq k)$ pour une unité quelconque, ainsi que la manière de mettre en œuvre la méthode générale de résolution de la problématique sous forme d'un algorithme.

Détermination de « p » pour un unité quelconque

Nous appelons « p » la probabilité théorique qu'une distance mesurée soit un multiple entier (plus ou moins une précision « Δe ») d'une unité quelconque « u ».

Calculer p revient donc à diviser :
→ le nombre de distances mesurables répondant au critère « est un multiple d'un entier à Δe près »,
→ par le nombre de distances mesurables total.

Calcul du nombre de distances mesurables total :
On l'obtient en divisant la plus grande distance mesurée « d_{max} » par la précision utilisée lors de la mesure des distances « Δd ».
Exemple : si nous mesurons des distances allant jusqu'à 1 200 mètres avec une précision de 0,1 mètre, nous pourrons mesurer 1 200 / 0,1 = 12 000 distances différentes.

Calcul du nombre de distances mesurables satisfaisant le critère cité plus haut :
Pour cela, on peut utiliser un algorithme qui va tester une à une toutes les distances mesurables. Pour chaque distance il va :
→ diviser la distance par u ;
→ tester si le résultat appartient à $[e - \Delta e \,;\, e + \Delta e]$ avec « e » un entier quelconque ;
→ ajouter 1 à une variable (initialisée à 0 au début) si le test est positif.
A la fin de l'algorithme, cette variable contiendra bien le nombre « d_{ok} » de mesures satisfaisant le critère « est un multiple d'un entier à Δe près ».

Exemple : si nous mesurons des distances allant jusqu'à 100 mètres avec une précision de 0,1 mètre, et que nous cherchons les multiples de l'unité 5 pour Δe = 0,01 l'algorithme va procéder ainsi :

- → *d_{ok} prend la valeur 0 ;*
- → *0,1/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, d_{ok} augmente de 1 ;*
- → *0,2/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, d_{ok} augmente de 1 ;*
- → *0,3/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, d_{ok} augmente de 1 ;*
- …
- → *99,9/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, d_{ok} augmente de 1 ;*
- → *100/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, d_{ok} augmente de 1 ;*
- → *L'algorithme retourne la valeur de d_{ok}.*

Calcul de p :

Une fois les deux paramètres déterminés, il suffit de les diviser :

$$p = \frac{d_{ok} . \Delta d}{d_{max}}$$

Détermination du paramètre « k »

k est le nombre de distances, parmi les distances mesurées, qui sont des multiples entiers à plus ou moins Δe de l'unité u testée. Il peut être déterminé avec un algorithme identique à celui que nous utilisons pour trouver d_{ok}.

L'algorithme va tester une à une toutes les distances mesurées. Pour chaque distance il va :

- → diviser la distance par u ;
- → tester si le résultat appartient à [e – Δe ; e + Δe] avec « e » un entier quelconque ;
- → ajouter 1 à k (initialisée à 0 au début) si le test est positif.

Exemple : si nous avons mesuré 6 distances (d_a à d_f) avec une précision de 0,1 mètre, et que nous cherchons les multiples de l'unité 5 pour Δe = 0,01 l'algorithme va procéder ainsi :

- → *k prend la valeur 0 ;*
- → *d_a/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *d_b/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *d_c/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *d_d/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *d_e/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *d_f/5 appartient-il à [e – 0,01 ; e + 0,01] avec e un entier ? Si oui, k augmente de 1 ;*
- → *L'algorithme retourne la valeur de k.*

Détermination de $P(X \geq k)$

X est une variable aléatoire représentant le nombre de distances mesurées exprimables en nombres entiers plus ou moins Δe dans l'unité u testée.

$P(X \geq k)$ est donc la probabilité, pour une unité testée, de trouver au moins k distances mesurées exprimables en nombre entiers plus ou moins Δe dans cette unité, parmi toutes les distances mesurées. Si $P(X \geq k)$ est très faible, c'est que le hasard peut difficilement être le seul responsable de la récurrente apparition de l'unité testée.

$P(X \geq k)$ se calcule en utilisant une loi binomiale car le fait qu'une distance mesurée soit ou non exprimable dans une unité n'influence pas la probabilité que la distance mesurée suivante soit exprimable dans cette même unité : les tests sont donc indépendants les uns des autres.

La méthode d'utilisation de la loi binomiale est décrite dans le document « *Méthode d'étude de la probabilité d'apparition d'angles remarquables dans un complexe mégalithique* », pages 3 et 4. Nous n'y reviendrons donc pas et nous invitons les personnes ne connaissant pas cette méthode à se reporter au document cité. Le résultat que nous y démontrons est le suivant :

$$P(X \geq k) = 1 - \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{(n-i)}$$

Dans notre cas, les paramètres p et k sont ceux que nous avons évoqué dans les parties précédentes, et n représente le nombre de tests réalisés, c'est à dire le nombre de distances mesurées, que nous pouvons noter « n_d ». Nous avons donc :

$$P(X \geq k) = 1 - \sum_{i=0}^{k} \binom{n_d}{i} \left(\frac{d_{ok} . \Delta d}{d_{max}}\right)^i \left(1 - \frac{d_{ok} . \Delta d}{d_{max}}\right)^{(n_d - i)}$$

Mise en œuvre de la méthode générale via un algorithme

Puisque les calculs présentés jusque là doivent être faits pour toutes les unités que l'on souhaite tester, ils doivent être répétés un très grand nombre de fois. La meilleure méthode est donc de créer un algorithme qui va automatiquement les effectuer pour une plage d'unité $[u_{min} ; u_{max}]$ que nous lui aurons précisée. Il faudra aussi préciser le pas « Δu » des unités à tester. Voici une structure générale de cet algorithme :

Entrées :

Saisir Δe ;
Saisir Δd ;
Saisir d_a, d_b, d_c... d_{nd} ;
Saisir u_{min} ;
Saisir u_{max} ;
Saisir Δu ;

Initialisations :

u = u_{min} ;
i = 0 ;

Traitement :

Supprimer les résultats des précédents calculs
Tant que (u ≤ u_{max})
Afficher u en [i ; 0] ;
Calculer p et l'afficher en [i ; 1] ; *// utiliser l'algorithme présenté pages 1 et 2*
Calculer k et l'afficher en [i ; 2] ; *// utiliser l'algorithme présenté page 2*
Calculer P(X ≥ k) et l'afficher en [i ; 3] ; *// utiliser le calcul présenté page 3*
i = i +1 ;
u = u + Δu ;
Fin Tant que
Trier les u dans l'ordre des P(X=k) correspondants croissants dans un nouveau tableau

En triant les u dans l'ordre proposé en fin l'algorithme, on fait remonter les unités dont les récurrentes utilisations sont les plus éloignées de ce qu'on aurait pu trouver dans une implantation hasardeuse des mégalithes du complexe étudié.