Chapitre 3.1

Introduction : « qui ? »

Anna DEGIOANNI

Aix Marseille Univ, CNRS, Minist Culture, LAMPEA, Aix-en-Provence

Dans cette partie, nous allons aborder la question : *qui ?*

Qui a laissé les traces étudiées par les archéologues ? La réponse est intuitive et simple : des êtres vivants, soit des humains soit d'autres animaux.

Concernant les humains, il y a une discipline qui les étudie de façon privilégiée : l'anthropologie (autrefois dite *anthropologie physique* ; elle s'appelle aujourd'hui *anthropologie biologique*).

Le début de cette discipline au XIX° siècle est strictement liée aux mathématiques et aux statistiques. On peut presque parler d'anthropométrie[1] pour les premiers 50 ans de cette discipline.

Un grand nombre d'outils statistiques, dont beaucoup font encore partie du répertoire standard des sciences, ont été conçus pour répondre aux questions des anthropologues.

Par exemple, le coefficient de corrélation de Pearson (*cf* 5) a été créé pour comprendre les relations existantes entre la taille des parents et celle des enfants ; l'on peut également citer les études portant sur les dimensions relatives des os du squelette (par exemple, la dimension du fémur par rapport à la stature) etc.

A partir des année 50, l'anthropométrie a connu une augmentation d'application biomédicales avec des travaux cliniques et expérimentaux. Il s'agissait de construire un plan expérimental avec la création de deux échantillons* qui subissent deux traitements différents et que l'on compare pour connaitre l'effet du traitement.

[1] Anthropométrie = mesure* de l'homme. Les premiers anthropologues ont pris des mesures de chaque partie de notre squelette et ont réalisé des comparaisons entre individus*.

A partir des année 70, le livre de chevet des anthropologues à travers le monde devient l'une des éditions du texte de Sokal, qui, avec son manuel de biométrie co-écrit avec F. J. Rohlf (Sokal & Rohlf, 2012) a profondément influencé la formation à l'analyse des données*, non seulement en anthropologie mais également en écologie et en biologie évolutive.

Entre temps, le monde de l'anthropologie se partage en deux grands groupes de recherche qui nécessitent des méthodes statistiques différentes :

- d'un côté, des anthropologues qui travaillent sur des données génétiques dont la quantité devient rapidement importante, et
- d'un autre coté, des anthropologues qui travaillent sur des données fossiles dont le nombre est très réduit.

Si certaines analyses sont toujours réalisables avec les méthodes existantes (comme les statistiques descriptives *cf* 4), la nature de certaines données nécessite la création de nouvelles méthodes d'analyses.

De plus en plus utilisé, le logiciel en libre-service R met à disposition de multiples packages prêts à l'emploi, adaptés aux questionnements des différentes disciplines ; ces packages ne nécessitent pas de connaissances particulières d'utilisation, mais nécessitent néanmoins d'en comprendre le résultat.

Dans les pages qui suivent seront présentés quatre exemples anthropologiques et un exemple paléontologique.

Le problème commun aux deux disciplines est la conservation au cours du temps des restes biologiques.

Naturellement, l'âge des fossiles et l'environnement dans lequel on les a retrouvés jouent énormément sur la qualité du matériel fossile (Brugal, 2017). Il reste néanmoins une question : le ou les fossiles issus des fouilles sont-ils des traces anecdotiques, ou bien peuvent ils nous renseigner sur d'autres vivants avec lesquels ils ont vécus ? La réponse est donnée par I. Séguy (*cf* 3.2).

E. Herrscher (*cf* 3.3) nous propose de réfléchir sur l'attitude à avoir lorsque l'on a à disposition des groupes de taille réduite. Le premier réflexe, qui n'est pas forcément le meilleur, est de tout regrouper pour permettre la mise en place de statistiques « robustes ». Il est néanmoins judicieux de vérifier *via* des tests spécifiques pour les petits échantillons si cela est possible.

G. Goude et K. Gerdau-Radonić (*cf* 3.4) nous expliquent et justifient l'utilisation de méthodes « ad hoc » pour des questionnements très particuliers.

Séguy et al (*cf* 3.5) donnent un exemple d'application de la méthode d'inférence bayésienne pour une problématique cruciale dans les études archéologiques : comment estimer la structure par âges d'une population à partir d'un faible nombre de squelettes archéologiques.

Le dernier exemple de Brugal *et al.* (*cf* 3.6) nous emmène dans le domaine de la paléontologie, discipline qui excelle dans la création et l'utilisation d'outils statistiques afin de mieux comprendre/connaitre la classification* des taxons ; la classification ascendante hiérarchique y est par exemple détaillée. Un exemple est aussi donné concernant l'analyses des traces laissées sur les ossements.

Naturellement, ces exemples ne sont pas exhaustifs de l'utilisation des méthodes statistiques pour répondre à la question « *qui ?* », mais ils proposent un panorama assez complet des méthodes et des problématiques actuelles.