

Quelle est la longueur des clous de 6,35cm ?

Mesures & incertitudes

There is no true value of anything.
W.E. Deming¹ [1900 – 1993]

Mesures & sources d'erreurs



FIG. 1 – L'arbalétrille, ou bâton de Jacob (gravure de Johannes Van Keulen, 1760), était utilisée pour obtenir la latitude d'une position sur la mer.

¹Tiré de sa préface au livre de [Walter Andrew Shewhart](#) [1891 – 1967], *Statistical Method from the Viewpoint of Quality Control*. (1938), réédité (1986) avec cette nouvelle préface par Dover Publications, New York NY.

Dès le XVI^e siècle, mesurer une position en mer est devenu un impératif pour ne pas se perdre et arriver à bon port lorsqu'on quitte les côtes avec leurs repères répertoriés.

L'[arbalestrille](#), ou bâton de Jacob, fut un des premiers instruments de mesure qui permettait de faire le point. L'opérateur de l'arbalestrille prenait l'angle entre, d'un côté du marteau coulissant (la partie transverse) fixé sur l'horizon, et de l'autre, soit le soleil à son zénith, ou midi, qu'il fallait savoir déterminer, la date aussi était importante, soit l'étoile polaire.

Pour obtenir cet angle, le mesureur devait déplacer le marteau coulissant le long de la règle graduée (l'autre partie), obtenir ainsi l'altitude en degrés du soleil (ou de l'étoile polaire) et utiliser des éphémérides pour la transformer en latitude². Cet instrument fut inventé grâce à des calculs trigonométriques (déjà toute une technologie!), probablement au début du XIV^e siècle par [Levi Ben Gerson](#) [1288–1344], qui est né et vécu en Provence et dans le Languedoc. L'arbalestrille connut des perfectionnements pendant des siècles et ses lointains descendants sont les [sextants](#) qui furent en usage jusque vers la fin du siècle dernier. On imagine sans peine l'imprécision de ces instruments, même dans leurs versions les plus évoluées.

Facteurs d'imprécision. La prise de mesure sur des bateaux constamment en mouvement, soumis aux aléas climatiques (il fallait au moins un ciel clair!), devait être très délicate : la visée en plein soleil, aveuglante, et la nuit pour l'étoile polaire, encore plus délicate ; l'opérateur pas toujours habile à saisir l'angle sur sa règle ; pour les mesures sur le soleil, comment être sûr de la précision obtenue pour la lecture du midi local sur des horloges à ressorts et balancier à l'origine, soumises aux aléas climatiques, roulis, tanguages, etc. ; l'instrument lui-même calibré imprécisément ; enfin, les éphémérides qu'il fallait utiliser pour convertir un angle en latitude étaient forcément imprécises, calculées à la main, à une certaine précision bien sûr et avec des théories pas très exactes. Bref, que de sources d'erreurs, de facteurs d'imprécision !

Mais on pourrait croire que les perfectionnements technologiques des systèmes de mesurage finiront par éliminer les erreurs, qu'on pourra un jour obtenir une vraie mesure, une mesure sans erreur aucune. Il n'en est évidemment rien.

On illustre, sur cet exemple, l'approche qu'il faut adopter pour tout examen des valeurs d'une mesure : quels sont les facteurs qui influent sur cette valeur ? Pour la précision, quels sont les facteurs qui peuvent jouer sur celle-ci ? C'est la question fondamentale dans toute investigation scientifique. Une mesure est une chaîne, chaque maillon une source d'erreurs.

Tout scientifique en action doit se poser ces questions. Il n'est en effet pas de science sans mesures, et donc sans systèmes de mesurage. La première constatation que doit faire un étudiant en Statistique est donc de se convaincre que la *variation* ou l'*incertitude* est omniprésente et inévitable. Il convient donc de développer des techniques pour la contrôler, la réduire. Que ce soit dans la prise de mesures, dans leur saisie ou dans leur exploitation, aucune certitude n'est

²Cela pour le cas du soleil seulement. L'utilisation de l'étoile polaire était de lecture directe.

possible finalement, la précision n'est pas de ce monde... L'une des raisons d'être des technologies statistiques est très exactement là !

L'expérience

Rien de tel qu'une petite expérience dans un contexte très simplifié pour se convaincre de la difficulté à obtenir des mesures exactes. En généralisant pour des cas plus complexes, on admettra qu'il est impossible d'obtenir de mesures exactes ou certaines.

Dans une classe, on a mesuré 5 clous de 6,35cm (ou 2,5in).³ Ces clous furent mesurés par 19 étudiants (dits 'opérateurs' dans ce contexte), qui ont utilisé deux règles métriques (dits 'instrument de mesure'), la première ayant été utilisée deux fois par chacun des opérateurs sur chacun des clous.

On se demande bien sûr quelle est la vraie longueur de chacun des clous.

Facteurs d'imprécision. On a en principe des clous de mêmes tailles. Les trois facteurs d'imprécision sont donc

1. les clous ou objets à mesurer ;
2. les règles ou appareils de mesure ;
3. les opérateurs de ces appareils.

Pour toute mesure, il faut savoir en déterminer les facteurs qui jouent sur sa valeur et ses erreurs. On dispose pour cela d'une technique, les diagrammes d'Ishikawa, ou diagramme des causes et des effets, qui est exposée en détails en classe, sur plusieurs exemples.

³Les dictionnaires annoncent que cette ancienne unité, *inch*, se traduit au Canada par 'pouce'. Sa valeur exacte est établie en termes du SI, comme la plupart de unités de mesure de l'ancien système *British Imperial* et du *United States Customary measure*, et vaut *exactement* 2,54cm depuis 1959. Rappelons que cet ancien système de mesures, en cette ère de globalisation, est encore en vigueur dans seulement trois pays : le Libéria, fondé au début du XIX^e siècle par des esclaves étatsuniens libérés (on dit aussi affranchis) qui, sous l'égide de la Société américaine de colonisation (sic), ont réintégré l'Afrique au grand dam des populations autochtones ; le Myanmar, anciennement la Birmanie ; et enfin les États-Unis. Tous pays qui ne brillent pas particulièrement par leur progressisme. Le Canada, de par sa proximité avec les États Unis, fonctionne dans une sorte de système commercialement mixte quoique pas entièrement étatsunien mais mâtiné du système *British Imperial*. C'est ainsi, par exemple, qu'on trouve des contenants de jus de 1,89l qui valent bien sûr 2 *american quarts* liquides, à ne pas confondre avec les *quarts* de matière sèche, toujours dans le système étatsunien, qui sont légèrement plus grands (1 101,22cm³, au lieu de 946,35ml pour les *liquid quarts*), ou encore avec les *quarts* dans le système Impérial (1 136,52cm³) où n'existe qu'une seule sorte de *quart*, qui doivent bien exister commercialement ailleurs au Canada ; de même une grande partie des bouteilles de bière de contenance 355ml qu'on trouve partout au Canada, qui n'est autre que 12 *fl.oz* étatsunienne, nombre bien plus simple, n'est-il pas... Quoi qu'il en soit, certains de ces nombres doivent être arrondis. Sont-ce les valeurs en *fl.oz* ou les valeurs en SI ? On pourra consulter le passionnant livre de Ken Adler, *Mesurer le monde. 1792-1799 : l'incroyable invention du mètre*. Paris : Flammarion, 2005. Notamment pour les enjeux scientifiques, politiques & commerciaux.

Protocole expérimental. Il importe que les conditions du déroulement de l'expérience soient planifiées avec soin. Au besoin, on commencera par une expérience pilote afin de s'assurer au préalable que des détails n'ont pas été oubliés. A posteriori, il n'y a plus qu'à éventuellement recommencer toute une expérience mal conçue, si elle s'avère impossible à réparer en cours de route... Il convient aussi de déterminer le gabarit pour la saisie des données qu'auront prises et notées les opérateurs. C'est le plus souvent dans un tableur électronique. Ce gabarit est déterminé en pensant aux objectifs de l'expérience.

1. *Utilisation de l'instrument.* On a au préalable indiqué aux opérateurs le fonctionnement de l'appareil de mesure. Dans notre cas, on place la règle au hasard le long du clou (et non avec le zéro d'un côté), ce qui permet d'avoir une précision plus grande. En effet, le mesureur (qui doit savoir additionner...) dispose de deux fractions de millimètres, une de chaque côté du clou, pour déterminer la longueur du clou.
2. On choisit au hasard un certain nombre de clous — de préférence impair, car alors les mesures ont un 'centre' naturel, i.e. une médiane —, dans la même boîte de clous.⁴
3. On se donne suffisamment de temps entre chaque séance de prises de mesure pour éviter de les rendre dépendantes l'une de l'autre pour chaque opérateur. Pour chaque séance de mesures, on présente les objets à mesurer dans un autre ordre. Ce qui permet d'éviter les interférences dues à l'ordre des objets à mesurer. Il faut cependant tenir le catalogue des ordres pour regrouper sans erreur les mesures de chaque clou.
4. On présente les objets à mesurer (les 5 clous) collés sur un carton avec la règle (instrument de mesure), et une feuille de papier pour saisir les mesures obtenues, successivement à chacun des opérateurs (mesureurs). Il importe que les mesures des opérateurs ne soient pas connues aux autres mesureurs, afin de s'assurer que les mesureurs soient le plus possible indépendants les uns des autres. (Cliquer sur l'icône à droite pour quelques illustrations.)
5. L'opérateur saisit sur la feuille de saisie des données, les mesures qu'il a déterminées, et signe la feuille. Il est important de connaître l'identité des opérateurs tout simplement par le principe de *traçabilité* : il faut retrouver les sources d'erreurs. Ce principe joue toujours dans les systèmes de production.
6. Le gabarit de saisie des données ayant été déterminé au préalable, on saisit les données, tout en prenant des précautions pour la vérification de la saisie. Des erreurs de transcriptions sont toujours possibles, et, dès ce stade, on peut identifier et corriger des données aberrantes.



⁴On aurait pu aussi, dans une perspective de contrôle de qualité de la production, choisir des clous au hasard dans des boîtes différentes. Ici, pour montrer les imprécisions de notre système de mesure (de n'importe lequel à cet égard), on essaie d'avoir les clous les plus homogènes possible. Parfois, donc, au contraire on désire maximiser la variabilité des produits à mesurer.

Traitement des données

Objectifs. Quels sont les objectifs de notre expérience ? Avant tout, déterminer si les propriétés du système de mesurage permettent de déterminer correctement la taille des clous. On se place en réalité dans un contexte de *métrologie*, fréquent dans les contextes industriels.

Ce n'est qu'indirectement que nous en déduirons si oui ou non on peut trouver de « *true value* » pour les longueurs des de chacun des clous.

Il va sans dire que la métrologie est un chapitre important, essentiellement de nature statistique, des sciences de l'ingénieur, et que nous ne faisons ici qu'en effleurer la surface.

En cliquant sur l'icône à droite, on pourra obtenir les données de l'expérience, ainsi que les premiers traitements des données.

La première feuille du fichier Excel comporte les données saisies de façon 'naturelle'. À dire vrai, avec quelques éliminations pour ne conserver que des données complètes.

L'expérience a utilisé deux instruments de mesure (règles), le premier ayant été utilisé deux fois par chaque opérateur et pour chacun des 5 clous, le second une seule fois. Normalement on aurait dû faire une autre séance de saisie avec le second instrument de mesure. Des raisons pratiques liées l'utilisation optimale du précieux temps des étudiants a quelque peu raccourci l'expérience. Enfin, nous n'avons pas conservé toutes les données, seules les données des sujets ayant participé aux trois parties de l'expérience ont été saisies sur le fichier. De la trentaine de sujets au départ, l'attrition naturelle⁵ a réduit notre échantillon d'opérateurs à 19...

Une première transformation des données a déjà été accomplie, une fois que la pensée s'y est vraiment appliquée... Au départ les données ont été saisies sur un gabarit près des feuilles de saisies par les opérateurs, on évite ainsi une source d'erreurs. Mais un peu de pensée aurait pu éviter cette étape.

On trouve les données transformées sous une forme plus près d'un fichier statistique de données, à savoir une matrice

sujets \times variables.

Les lignes sont les 'sujets' ou unités statistiques, les colonnes les 'variables' ou les descriptions des unités statistiques. Ces colonnes sont des caractéristiques des unités statistiques ou des mesures à proprement parler.

Cette première transformation permettrait un certain nombre d'observations, à savoir les moyennes des tailles mesurées des 5 clous par sujet et par instrument de mesure et par essai de mesure (on a deux essais pour la première règle et un seul pour la deuxième).

⁵ Absentéisme, toutes les raisons sont bonnes, c'est le début de l'année, etc. ; certains étudiants ne signaient pas leur fiche de saisie ; lectures illisibles.

Clous



Data

Mais ces moyennes ne font pas nécessairement de sens, encore faut-il savoir si les mesures diffèrent de fois en fois ; si les deux règles sont identiques, i.e. mesurent de la même façon ; si les clous sont ‘identiques’, ou de même taille.

1. (2 points) Pouvez-vous identifier une autre raison qui ferait que mettre toutes les mesures des clous dans un même sac et en prendre la taille moyenne ne serait pas légitime ?

On présentera tout au long du cours des technologies statistiques permettant d’explorer les réponses à ces questions. Pour l’instant on se contentera de constatations descriptives (et non à proprement parler statistiques), dont on tirera des conclusions reliées à nos objectifs.⁶

Une première permutation des colonnes de la matrice précédente donnée dans le fichier, suivi d’un tri croissant (dit personnalisé dans Excel) selon certaines colonnes, permet de mieux répondre aux questions suivantes. Cette nouvelle matrice est aussi donnée dans le fichier.

2. Examen des performances avec le premier instrument de mesure.

Tout d’abord, on estime la *qualité* des opérateurs par leurs stabilité dans le mesurage avec la première règle. On doit construire des matrices particulières avec, en lignes les opérateurs, et en colonnes des mesures de la qualité des mesures. Donc, on utilisera d’abord une mesure de dispersion et on a choisi l’écart type (on aurait pu choisir aussi l’écart entre la mesure maximum et la mesure minimum).

- (a) (3 points) Exprimer en mots pourquoi la qualité d’un opérateur peut être quantifiée par son écart type dans l’utilisation répétée d’un instrument sur un objet à mesurer. Pourquoi utiliser plusieurs objets à mesurer dans ces répétitions de mesurages. Une fois qu’on a réduit les opérateurs à ceux qui ont des qualités voisines, comment mesurer la qualité des instruments de mesure, s’ils sont nombreux ? Enfin, comment s’assurer que les objets à mesurer sont homogènes, i.e. de ‘mesures’ voisines ? Comment peut-on alors être certain de la mesure d’un objet ? (On ne s’attend pas à des réponses parfaitement complètes et compétentes.)
- (b) (2 points) Obtenir, pour le premier instrument de mesure, les écarts types par opérateur et par clou. Vous devez créer une nouvelle matrice de données avec en lignes les 19 opérateurs, et six colonnes,

⁶On doit souligner que tout analyste de données doit savoir utiliser les manipulations de données grâce aux facilités des logiciels à sa disposition pour en tirer de l’information. Tout analyste sait également que les logiciels ne sont pas tous également souples : certaines manipulations et productions de graphiques — fort utiles en exploration de données, il importe en effet avant toutes choses de *voir* les données —, simples dans un logiciel, peuvent devenir fastidieuse dans un autre.

représentant les écarts types des deux mesures du premier instrument pour les six clous par opérateurs. Vous ajouterez deux colonnes supplémentaires comportant, par opérateur, les moyennes des écarts types ainsi que les écarts types de ces 6 écarts types.

Représentez graphiquement pour les 19 opérateurs ces qualités de leurs mesures avec le premier instrument, d'abord pour quelques clous pris un à un, puis pour quelques paires, enfin pour tous les clous simultanément. Vous représenterez aussi par opérateur les moyennes de ces écarts types, de même que leur écart type.

Vous semble-t-il que les opérateurs ont des performances identiques, ou pouvez-vous faire l'hypothèse que quelques opérateurs sont un peu à part ? Justifiez en quelques mots vos jugements.

- (c) (2 points) Procédez de même avec les moyennes des deux mesures par opérateur et par clou. Vous avez à construire une matrice analogue à celle construite au numéro précédent (2b)

Représentez graphiquement cette suite de moyennes par opérateur, pour quelques clous pris un à un, comme au numéro précédent. Pouvez-vous encore identifier des opérateurs singuliers par rapport aux autres ? Vous semble-t-il que les clous sont tous vaguement identiques, ou si quelques-uns détonnent par rapport aux autres ? Vous justifiez en quelques mots vos jugements. Les graphiques que vous avez produits sont-ils tous également utiles ?

- (d) (2 points) Après avoir obtenu les valeurs moyennes, minimales et maximales sur tous les opérateurs des tailles des clous, et ce pour chacun des clous, et en comparant les graphiques des moyennes de tailles des clous par opérateur, pour quelques paires de clous puis pour quelques triplets, enfin pour l'ensemble, pour le premier instrument toujours, vous apparaît-il que certains opérateurs se distinguent des autres ? Vous justifiez en quelques mots vos jugements.
- (e) (2 points) Recommencez en partie les graphiques plus haut en ayant supprimé ces opérateurs singuliers. Sauriez-vous dire si les clous sont à peu près identiques ? Justifiez en quelques mots. Rédigez vos conclusions à la suite de ces considérations.

3. Performances comparées des deux instruments de mesure. Nous introduisons également des mesures statistiques de dispersion des données. Nous conservons dans cette étape l'ensemble des opérateurs.

Construire maintenant, par des colonnes en sus de la matrice précédente, les rangées des mesures pour chaque clou du deuxième instrument de mesure. Les rangées de cette matrice sont les mêmes que pour la matrice précédemment construite.

Ajoutez les éléments suivants en rangées : les minimum, maximum, moyenne & médiane des valeurs pour le premier et deuxième instrument pour les 19 opérateurs, et ce pour chacun des clous. Ce sont là des indicateurs de localisation des données. De même, vous ferez calculer les rangées suivantes

qui sont des indicateurs de dispersion des données : les différences ‘max - min’, les écarts types et les erreurs types, soit les valeurs des écarts types divisés par \sqrt{n} où n est le nombre d’observations pour les opérateurs. Ici $n = 19$.

On se retrouve avec les observations pour chacun des 19 sujets (en lignes ou rangées de la matrice obtenue) et pour chacun des 6 clous, des valeurs de la moyenne des longueurs observées des deux mesures pour le premier instrument, et de la valeur observée de la taille du clou ; avec en plus pour chaque clou et instrument, les indicateurs de la localisation et ceux de la dispersion des tailles des clous, soit 12 colonnes en tout.

- (a) (2 points) Obtenez les graphiques qui comparent pour un certain nombre de clous, les 19 mesures des opérateurs avec les premier et second instrument de mesure. Sauriez-vous formuler une hypothèse sur les propriétés des deux instruments de mesure ? Justifiez votre réponse.
 - (b) (2 points) Les opérateurs qui apparaissent singuliers pour le premier instrument le sont-ils encore pour le deuxième ? Justifiez votre réponse.
4. Suite de la comparaison des deux instruments de mesure. Pour cette question, vous devrez tamiser les données en enlevant les opérateurs singuliers que vous avez identifié au numéro 3.

- (a) (4 points) Vous utiliserez maintenant les valeurs pour le premier instrument et pour chaque clou des moyennes pour les 19 opérateurs calculées au numéro 3, de même que les valeurs des erreurs types.

Calculez pour chaque clou un intervalle de largeur $2 \times$ ‘erreurs types’ centré sur la moyenne du premier instrument.

Cet intervalle est une bonne approximation de l’*intervalle de confiance de la moyenne* qui indique une région probable à environ 95% (cela sera expliquée en grands détails plus tard dans le cours) pour la localisation de la moyenne. On a un concept comme celui-ci dans tous les sondages : la moyenne indiquée sera valide avec cette marge d’erreur 19 fois sur 20 (ou 95%), i.e. sera située 19 fois sur 20 dans cet intervalle, dit de confiance à 95% pour un autre échantillon.

Examinez enfin, pour chaque clou, si les valeurs des tailles moyennes obtenues du deuxième instrument sont situées dans l’intervalle dit de confiance.

Vous conclurez sur les propriétés du deuxième instrument de mesure relativement au premier.

- (b) (4 points) Obtenez maintenant le graphique en fonction des 5 clous des valeurs moyennes pour les premier & second instrument. Utilisez mentalement les intervalles de confiance pour le premier instrument de mesure, en les représentant mentalement autour des moyennes pour chaque clou. Que constatez-vous sur les propriétés des deux

instruments de mesure ? Cela est-il confirmé par les localisations des 5 moyennes pour chacun des deux instruments de mesure ?

5. Vous disposez maintenant de 10 mesures pour les moyennes des longueurs des clous, 5 pour chaque instrument de mesure. Ce qui reste des opérateurs vous est apparu assez homogène dans leurs mesures pour ce regroupement moyen des mesures de chacun des 5 clous.
 - (a) (3 points) Calculez maintenant, pour chaque instrument de mesure à partir de ces 5 observations des tailles moyennes, leurs moyennes (ce sont alors des moyennes de moyennes) et erreurs types. Déterminez les deux intervalles de confiance de ces moyennes, un pour chaque instrument.

Le nominal de la longueur des clous est de 6,35cm. Est-ce que cette valeur nominale est à l'intérieur de deux intervalles de confiance. Seriez-vous prêts à penser que la valeur nominale est bien celle annoncée ?
 - (b) (2 points) Quelle est la valeur exacte de la longueur du troisième clou ?
-

Nous voici arrivé à bon port ! Non, la valeur de la longueur d'un clou ne peut être mesurée avec précision. Non mesurable, la valeur exacte des tailles ne peut donc exister... Enfin, 'exister' dans un sens que nous ne préciserons pas... Mais alors qu'est-ce qu'on trouve dans le fin fond de la réalité naturelle, celle des choses matérielles ? Nous ne répondrons pas plus à cette question métaphysique !

Inutile, enfin espérons-le, de faire une démonstration générale de la proposition « *There is no true value of anything* » (Deming). Ce qui ne veut pas dire que nous ne convergerons jamais vers une valeur exacte, « réelle », mais parions que non ! On pourrait tenter des expériences sur les mêmes principes que ceux illustrés ici pour des instruments de navigation de la fin du XX^e siècle (Fig. 3 & 4), plutôt que pour la longueur des clous, on verrait encore des sources d'erreurs et des mesures imprécises ! On se rapproche des mesures exactes, on ne les connaîtra jamais.



FIG. 2 – Instruments de navigation en usage fin du XVIII^e siècle. Planche tirée de l'Encyclopédie de d'Alembert & Diderot.

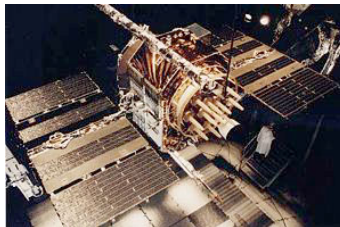


FIG. 3 – Un satellite NavStar appartenant à la constellation du GPS (fin du XX^e siècle).



FIG. 4 – À gauche : des (récepteurs) GPS de poche ; à droite : un GPS de bateau qui permet de déterminer et contrôler automatiquement des trajectoires.

Ces figures donnent quelque peu le vertige...