

Allométrie pour les nuls

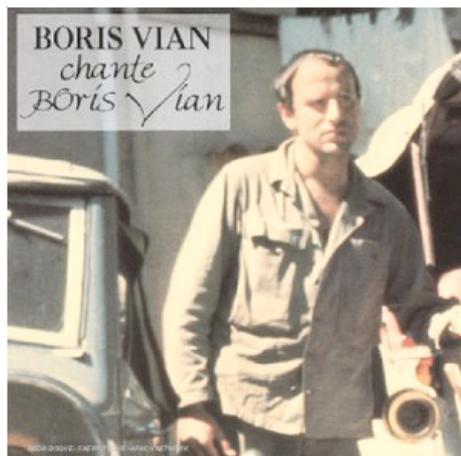
P^r Jean R. LOBRY

Ce document est une introduction tout en douceur aux concepts de l'allométrie avec plusieurs niveaux de lecture. Il y a beaucoup de latin (du code ). Passez outre en première approche ! Umberto Eco a écrit dans l'Apostille au Nom de la rose *qu'un jeune garçon de dix-sept ans lui a dit qu'il n'avait rien compris aux discussions théologiques mais qu'elles agissaient comme des prolongements du labyrinthe spatial*. Par rapport au latin, les exemples de  sont **reproductibles**.

Table des matières

1	Vous n'êtes pas si nuls...	2
2	Vous êtes très très nuls...	7
3	Vers un compromis	9
3.1	Que nous dit la valeur de la pente?	12
3.2	Que nous disent les écarts à la droite?	16
4	Quelques pièges de l'allométrie	17
4.1	Le problème des points expérimentaux	17
4.2	Le problème du tout et de la partie	17
4.3	Le problème de Galton	18
4.4	Le problème de l'interprétation	19

1 Vous n'êtes pas si nuls...



Quand j'aurai du vent dans mon crâne
 Quand j'aurai du vert sur mes osse
 P'tête qu'on croira que je ricane
 Mais ça sera une impression fosse
 Car il me manquera
 Mon élément plastique
 Plastique tique tique
 Qu'auront bouffé les rats
 [...]

Et puis je n'aurais plus
 Ce phosphore un peu mou
 Cerveau qui me servit
 À me prévoir sans vie
 [...]

Vian, B. (1920-1959)
 Quand j'aurai du vent dans mon crâne.
 Crédit photo : Disque B0000084FK.

Cette première partie est une traduction et adaptation (assez libre) de la page de Jim MOORE¹, anthropologue à l'Université de Californie à San Diego². C'est un exemple d'application de l'allométrie aux mammifères, dont vous faites incidemment parti en tant que membre de l'espèce *Homo sapiens sapiens*. Cette approche allométrique est très puissante pour comprendre le fondement des études comparatives ; même si elle peut sembler quelque peu obtuse à première vue, quelques exemples devraient vous aider à en cerner l'intérêt. A cause du lien [éventuel] entre la taille du cerveau, votre propre nullité, et l'évolution humaine, les exemples ci-dessous concernent tous la taille relative du cerveau, mais notez bien que la méthode est applicable à l'étude de bien d'autres caractères morphologiques, et ce à différents niveaux (individu, population, espèce, etc).

La plupart des caractères biologiques augmentent comme une fonction puissance de la taille du corps. Par exemple, il faut plus d'énergie pour faire vivre un éléphant qu'une souris. Plus la taille de l'organisme augmente, plus il lui faut d'énergie pour survivre. Côté cognitif, si les corps sont contrôlés par les cerveaux, il est raisonnable de penser que plus le corps est grand, plus le cerveau qui le pilote devra être grand (plus de nerfs pour coordonner l'action de plus de muscles, etc). Cependant, de telles dépendances vis-à-vis de la taille du corps sont rarement dans un rapport 1:1 (c'est-à-dire qu'il est rare qu'un accroissement d'une unité de la taille du corps produise exactement un accroissement d'une unité du métabolisme, de la taille du cerveau, ou de quoi que ce soit).

Donc, nous nous intéressons à la relation entre la taille du corps et du cerveau, nous nous attendons à ce que les gros animaux aient de plus gros cerveaux, et nous aimerions en savoir plus sur cette relation et ce pour *deux raisons* :

1. savoir quelque chose sur la relation générale entre la taille du corps et du cerveau peut, éventuellement, nous permettre de dire des choses intéressantes sur l'intelligence en général (votre degré de nullité par exemple).
2. s'il y a une relation générale, alors nous voulons la neutraliser lorsque nous discutons de la taille du cerveau en relation avec la cognition : les vaches

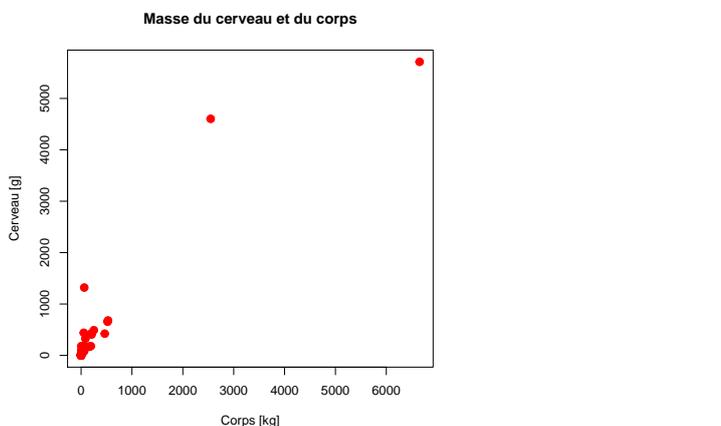
1. <http://weber.ucsd.edu/~jmoore/courses/allometry/allometry.html>

2. <http://www.ucsd.edu/>

ont un plus gros cerveau que les singes, mais cela est sans doute plus lié au fait que les vaches sont de **grosses** vaches qu'à leur niveau d'intelligence intrinsèque.

Pour commencer, représentons graphiquement la relation entre la taille du cerveau et du corps pour 62 espèces de mammifères. Le jeu de données [1] utilisé est extrait de Allison et Cicchetti (1976), et, chose extrêmement rare dans le domaine, disponible pour l'ensemble de la communauté scientifique sur le web. Il a été intégré dans le jeu de données `mammals` de la bibliothèque standard MASS sous .

```
library(MASS)
data(mammals)
attach(mammals)
plot(x = body, y = brain, pch = 20, cex = 2, col = "red",
     xlab = "Corps [kg]",
     ylab = "Cerveau [g]",
     main = "Masse du cerveau et du corps")
```

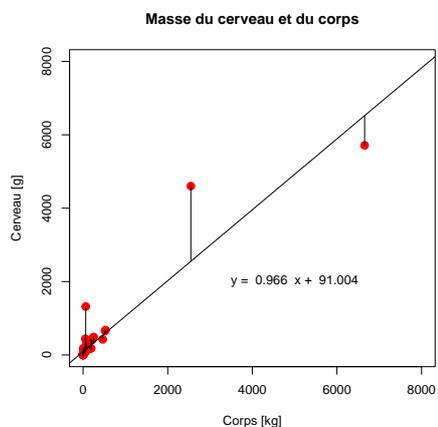


Voilà donc à quoi ressemblent les données. Notez que les masses corporelles sont en kilo (1000 g), il faut donc multiplier la masse des corps par 1000 pour les avoir dans la même unité que celle de la masse des cerveaux. Assurément, les espèces ayant un gros corps ont tendance à avoir un gros cerveau (le scoop!). Mais comment pourrions nous exprimer cette relation ? Pourquoi ne pas essayer de faire une régression de la taille du cerveau sur celle du corps ? Il y a plusieurs façons de s'y prendre, mais le principe reste le même : c'est une droite qui s'ajuste « le mieux » possible aux observations (si nous arrivons à nous entendre sur ce que « le mieux » veut dire). La droite elle-même peut être décrite par une simple équation de la forme :

$$y = ax + b \tag{1}$$

où y est la masse du cerveau prédite, x la masse corporelle observée, a la pente, et b la valeur pour laquelle la droite coupe l'axe vertical.

```
lm1 <- lm(brain~body)
plot(x = body, y = brain, pch = 20, cex = 2, col = "red",
     xlim = c(0, 8000), ylim = c(0, 8000),
     xlab = "Corps [kg]",
     ylab = "Cerveau [g]",
     main = "Masse du cerveau et du corps")
abline(coef = lm1$coef)
segments(body, brain, body, predict(lm(brain~body)))
text(5000, 2000, paste("y = ", round(lm1$coef[2], 3), " x + ", round(lm1$coef[1], 3)))
```



```

sml1 <- summary(lm1)
sml1
Call:
lm(formula = brain ~ body)
Residuals:
    Min       1Q   Median       3Q      Max
-810.07  -88.52  -79.64  -13.02  2050.33

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  91.00440   43.55258     2.09  0.0409 *
body         0.96650    0.04766    20.28 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 334.7 on 60 degrees of freedom
Multiple R-squared:  0.8727,    Adjusted R-squared:  0.8705
F-statistic: 411.2 on 1 and 60 DF,  p-value: < 2.2e-16

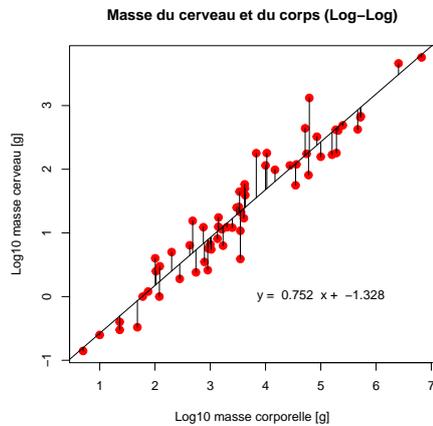
```

La droite de régression explique ici environ 87.3 % de la variance (r^2), autrement dit environ 87.3 % de la variation de la masse du cerveau, y , est prise en compte par la variation de la masse corporelle, x . Mais il y a quelques petits problèmes... En regardant la droite de régression, il est clair que l'on donne beaucoup trop de poids (c'est le cas de le dire) aux deux derniers points (l'éléphant d'Asie et l'éléphant d'Afrique). Toutes les autres espèces sont compactées en bas et à gauche du graphique, et il est extrêmement difficile d'avoir une idée visuelle des variations en leur sein. Que faire ? Nous allons convertir les données initiales en logarithmes de base 10 avant de les représenter, comme dans la figure ci-après (*nb* il est plus commode d'avoir les mêmes unités, la masse corporelle sera maintenant exprimée en grammes).

```

x <- log10(1000*body)
y <- log10(brain)
plot(x = x, y = y, pch = 20, cex = 2, col = "red",
      xlab = "Log10 masse corporelle [g]",
      ylab = "Log10 masse cerveau [g]",
      main = "Masse du cerveau et du corps (Log-Log)")
lm2 <- lm(y~x)
abline(coef = lm2$coef)
segments(x, y, x, predict(lm2))
text(5,0,paste("y = ",round(lm2$coef[2],3)," x + ",round(lm2$coef[1],3)))

```



```

slm2 <- summary(lm2)
slm2
Call:
lm(formula = y ~ x)
Residuals:
    Min       1Q   Median       3Q      Max
-0.74503 -0.21380 -0.02676  0.18934  0.84613

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.32793    0.10888  -12.20  <2e-16 ***
x             0.75169    0.02846   26.41  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3015 on 60 degrees of freedom
Multiple R-squared:  0.9208,    Adjusted R-squared:  0.9195
F-statistic: 697.4 on 1 and 60 DF,  p-value: < 2.2e-16

```

Comme vous pouvez le constater, quand les données sont exprimées en logarithmes, elles s'ajustent bien à une relation linéaire, et les points sont maintenant bien répartis, de sorte que nous pouvons bien mieux saisir ce qui se passe.

Maintenant : l'équation de la droite est de la forme habituelle $y = ax + b$, mais y et x représentent en fait ici **Log**(masse du cerveau) et **Log**(masse corporelle).

C'est une *équation allométrique*; « allo » vient du grec allos = « autre », dans ce cas « autre que métrique », c'est à dire non-linéaire (le contraire serait une équation linéaire ou isométrique, par exemple pour chaque kilo de masse corporelle ajouter 20 g de cerveau). Nous l'avons linéarisée par la double transformation logarithmique pour mieux la visualiser et nous faciliter le travail. Pour exprimer la même relation non-linéaire directement, c'est-à-dire avec une équation où x représente la masse corporelle, et non **Log**(masse corporelle), nous convertissons l'équation ci-dessous :

$$\text{Log}(y) = a[\text{Log}(x)] + \text{Log}(b) \quad (2)$$

dans la forme standard d'une équation d'allométrie :

$$y = bx^a \quad (3)$$

EXEMPLE : Quelle masse devrait faire le cerveau d'un rat géant africain, de masse corporelle égale à 1000 g, en se basant sur les régressions ci-dessus ?

1. Avec la forme $y = ax + b$ (donnée dans le graphe précédent), $y = 0.752x - 1.328$. C'est la relation *empirique* pour l'équation de la droite de régression de la masse du cerveau par la masse corporelle dans notre échantillon, et donc notre meilleur outil pour cette prédiction.

(mmm... , c'est quoi le Log de 1000, déjà ?)

$\text{Log}(1000) = 3$ (rappel : $1000 = 10^3$) donc

$\text{Log}(y) = 0.752 * 3 - 1.328 = 0.928$

$y = 10^{0.928} = 8.5$ g

2. Ou encore en utilisant $y = bx^a$, après avoir converti le b du Log :

$10^{-1.328} = 0.047$

$y = 0.047 * 10000^{0.752}$

= 8.5 g

3. Ou bien directement sous R :

```
10^predict(lm2)[rownames(mammals)=="African giant pouched rat"]
      13
8.45526
```

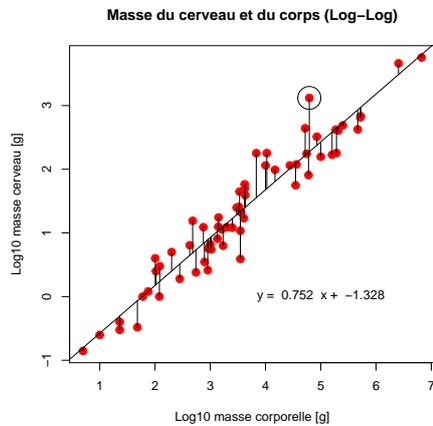
A comparer à la valeur observée de 6.6 g pour le rat géant africain. Il n'a donc qu'un cerveau de seulement 6.6 g alors qu'il devrait faire 8.455 g pour suivre la tendance générale.

Et pour vous, *Homo sapiens sapiens*, alors ?

```
mammals[rownames(mammals) == "Human", ]
      body brain
Human   62  1320
10^predict(lm2)[rownames(mammals) == "Human"]
      32
188.1233
```

Pour votre masse corporelle, vous devriez avoir un cerveau de 188 g, alors qu'il pèse ses bons 1320 g. Conclusion, selon l'équation d'allométrie au sein des mammifères, vous ne pouvez pas être aussi nuls que ça. Visualisons cela :

```
x <- log10(1000*body)
y <- log10(brain)
plot(x = x, y = y, pch = 20, cex = 2, col = "red",
      xlab = "Log10 masse corporelle [g]",
      ylab = "Log10 masse cerveau [g]",
      main = "Masse du cerveau et du corps (Log-Log)")
lm2 <- lm(y~x)
abline(coef = lm2$coef)
segments(x, y, x, predict(lm2))
text(5,0,paste("y = ",round(lm2$coef[2],3)," x + ",round(lm2$coef[1],3)))
points(x[rownames(mammals) == "Human"], y[rownames(mammals) == "Human"], cex = 4)
```



Voilà qui est fort satisfaisant pour votre ego, rapporté à la masse corporelle, vous avez le plus gros cerveau de tous les mammifères. Quels sont vos plus proches compétiteurs de ce point de vue ?

```
top10 <- order(-lm2$residuals)[1:10]
cbind(mammals[top10, ], 10^lm2$fitted.value[top10])
```

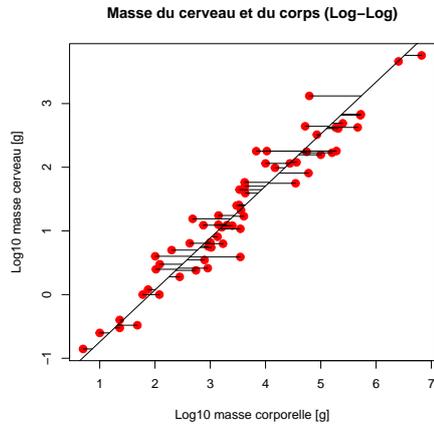
	body	brain	10^lm2\$fitted.value[top10]
Human	62.000	1320.0	188.123291
Rhesus monkey	6.800	179.0	35.720060
Baboon	10.550	179.5	49.692566
Owl monkey	0.480	15.5	4.869905
Chimpanzee	52.160	440.0	165.205861
Ground squirrel	0.101	4.0	1.509000
Patas monkey	10.000	115.0	47.732357
Verbet	4.190	58.0	24.821983
Arctic fox	3.385	44.5	21.144078
Red fox	4.235	50.4	25.022105

Des singes, on pouvait s'y attendre...

2 Vous êtes très très nuls...

Mais en fait, les choses ne se présentent pas ainsi. Le volume des neurones étant à peu près constant entre espèces, le nombre de neurones disponibles pour piloter un corps est proportionnel en première approximation à la masse du cerveau. La bonne question est donc : quelle masse corporelle peut-on espérer contrôler avec une unité centrale de telle capacité ? Pour apprécier l'optimalité d'une espèce, il convient donc de faire la régression de la masse corporelle sur la masse du cerveau, et non pas le contraire.

```
lm3 <- lm(x~y)
plot(x, y, pch = 20, cex = 2, col="red",
      xlab = "Log10 masse corporelle [g]",
      ylab = "Log10 masse cerveau [g]",
      main = "Masse du cerveau et du corps (Log-Log)")
abline(-lm3$coef[1]/lm3$coef[2], 1/lm3$coef[2])
segments(x,y,predict(lm3),y)
```

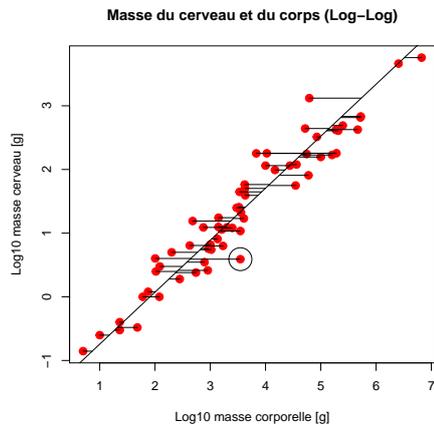


```
10^lm3$fitted.value[rownames(mammals) == "Human"]
32
540625
```

Autrement dit, avec l'unité centrale dont vous êtes dotés, vous devriez être capable de piloter une masse corporelle de 540 kg, alors que vous vous contentez d'un modeste 64 kg. C'est absolument et définitivement minable! Un gâchis considérable de ressources, vous êtes l'équivalent d'une machine à café pilotée par un supercalculateur. Pas de quoi pavoiser...

```
best <- order(-lm3$residuals)[1]
mammals[best,]
      body brain
Water opossum 3.5 3.9

lm3 <- lm(x~y)
plot(x, y, pch = 20, cex = 2, col="red",
     xlab = "Log10 masse corporelle [g]",
     ylab = "Log10 masse cerveau [g]",
     main = "Masse du cerveau et du corps (Log-Log)")
abline(-lm3$coef[1]/lm3$coef[2], 1/lm3$coef[2])
segments(x,y,predict(lm3),y)
points(x[rownames(mammals) == "Water opossum"],y[rownames(mammals) == "Water opossum"], cex = 4)
```





Chironectes minimus (l'opossum aquatique ou Yapok) est une petite merveille de programmation embarquée avec un cerveau de moins de 4 g pour contrôler un corps de 3.5 kg. Pour arriver à ce niveau de performance le cerveau humain devrait piloter un corps d'environ 5 tonnes. Crédit photo : http://animaldiversity.umz.umich.edu/site/resources/phil_myers/classic/chironectes_minimus.jpg

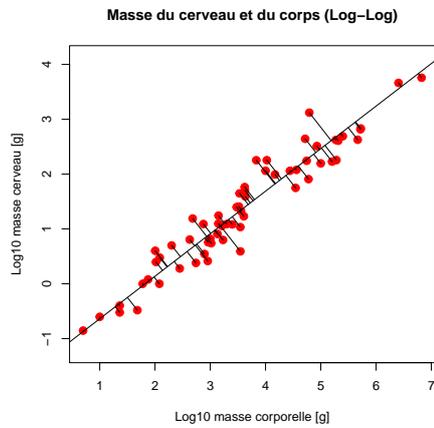
En tant que membre de l'espèce *Chironectes minimus*, je conteste avec la plus grande vigueur la validité du point de vue de la régression de la masse du cerveau sur celle du corps, c'est d'un anthropocentrisme des plus éhonté destiné uniquement à flatter l'ego des *Homo sapiens sapiens*.

3 Vers un compromis

Quel que soit le sens de la régression utilisée, cela revient à donner à une des deux variables étudiées le statut privilégié de variable de contrôle. Or, comme son nom l'indique, une variable de contrôle doit être contrôlable : l'expérimentateur doit pouvoir lui assigner une valeur selon son bon vouloir. Par exemple la température d'un milieu réactionnel est une variable de contrôle parce qu'elle est contrôlable avec un thermostat. Ici, il n'en est rien parce que l'on ne sait pas encore synthétiser *de novo* un mammifère fonctionnel ayant une masse corporelle donnée *a priori*. L'expérimentateur n'est ici en fait qu'un observateur qui se contente de mesurer les valeurs existant dans la nature. Il n'y a donc aucune raison valable pour privilégier une variable par rapport à l'autre, et il est préférable de faire une régression orthogonale³.

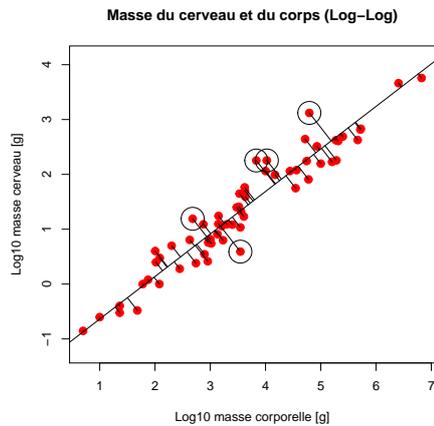
```
w <- as.data.frame(cbind(x,y))
plot(w,asp=1,pch=20, cex=2, col="red",
      xlab="Log10 masse corporelle [g]",
      ylab="Log10 masse cerveau [g]",
      main = "Masse du cerveau et du corps (Log-Log)")
cov <- var(w)
u <- eigen(cov, sym=TRUE)$vectors[,1]
p <- u[2]/u[1]
scal <- (x - mean(x)) * u[1] + (y - mean(y)) * u[2]
segments(x, y, mean(x) + scal * u[1], mean(y) + scal * u[2])
abline(c(mean(y) - p * mean(x), p))
```

3. Pour nos lecteurs de niveau avancé qui seraient tombé par hasard sur cette fiche en désespoir de cause : on trouvera dans le paquet  *smatr* de Warton [8] tous les outils pour aller plus loin dans l'analyse des relations allométriques [9].



Le critère d'ajustement utilisé ici consiste à minimiser la somme des carrés des distances entre les points observés et leur projection orthogonale sur la droite de régression, il n'y a pas une variable qui soit privilégiée par rapport à l'autre. Quelles sont les espèces les plus distantes de cette droite ?

```
res <- sqrt((x-mean(x) - scal*u[1])^2 + (y - mean(y) -scal*u[2])^2)
top5 <- order(-res)[1:5]
w <- as.data.frame(cbind(x,y))
plot(w,asp=1,pch=20, cex=2, col="red",
      xlab="Log10 masse corporelle [g]",
      ylab="Log10 masse cerveau [g]",
      main = "Masse du cerveau et du corps (Log-Log)")
cov <- var(w)
u <- eigen(cov, sym=TRUE)$vectors[,1]
p <- u[2]/u[1]
scal <- (x - mean(x) )*u[1] + (y - mean(y) )*u[2]
segments(x, y, mean(x) + scal*u[1], mean(y) +scal*u[2])
abline( c(mean(y) -p*mean(x), p))
points(x[top5],y[top5], cex=4)
mammals[top5, ]
      body brain
Human      62.00 1320.0
Water opossum  3.50   3.9
Rhesus monkey  6.80 179.0
Baboon      10.55 179.5
Owl monkey    0.48  15.5
```



On retrouve ainsi *Homo sapiens sapiens* et *Chironectes minimus* tout en haut du tableau. Ce sont les deux espèces qui s'éloignent le plus de la tendance

générale. *Homo sapiens sapiens* a une plus petite masse corporelle et un plus gros cerveau que la tendance générale, *Chironectes minimus* a une plus grande masse corporelle et un plus petit cerveau que la tendance générale.

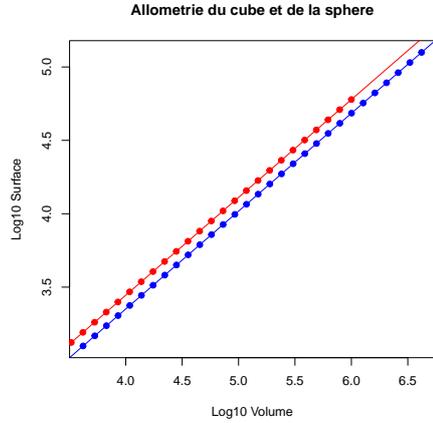
En résumé : la forme générale d'une équation d'allométrie est $y = bx^a$, où y présente la mesure en question, x est la taille (généralement la masse corporelle), a est l'exposant allométrique (qui nous donne la relation entre x et y), et b une constante (le coefficient allométrique). Par exemple, au sein des mammifères, le métabolisme basal (MB, en $Kcal.h^{-1}$) est donné par $MB = 3.8(\text{masse})^{0.75}$.

Ce qu'il faut bien comprendre c'est que dans cette équation (et dans les autres équations allométriques) les valeurs 3.8 et 0.75 sont obtenues de façon empirique. Vous avez pesé de nombreux animaux, mesuré leur MB, représenté les données sur un graphe en coordonnées log-log et trouvé que la droite de régression avait une pente de 0.75 et une ordonnée à l'origine de 3.8. Évidemment, il y aura des différences selon le jeu d'espèces que vous étudiez. Différents taxa auront des valeurs différentes pour b , et la valeur de b dépendra de ce que vous étudiez (ici MB). Cela signifie que vous pourrez trouver différentes valeurs pour, mettons, la pente de la droite résumant la relation entre la taille du cerveau et la masse corporelle : tout dépend de l'échantillon.

AAARRRGLLLL! Qu'est ce c'est que ce $B^{***}L$? Je veux dire : nous nous sommes ingurgités toutes ces stats pour découvrir que l'on obtient une réponse précise, mais des réponses différentes selon les échantillons? Quelle est cette arnaque? Qu'est ce que la pente nous *dit* que nous ne sachions pas déjà ???

Il appert que le métabolisme basal (MB) augmente environ comme une puissance 0.75 de la masse corporelle. Un babouin a besoin dans l'absolu de plus d'énergie qu'une marmotte, mais en consomme bien moins par gramme (pourquoi est-ce 0.75, on ne sait pas, c'est une observation empirique). D'un autre côté, pour des contraintes purement géométriques, la surface d'un objet de forme donnée augmente comme une puissance 0.67 de son volume (qui est quasiment assimilable à la masse tant que l'on reste dans le domaine animal). Nous pouvons illustrer ceci dans le cas d'une sphère et d'un cube :

```
vol.sphere <- fonction(r) { return(4/3*pi*r^3) }
surf.sphere <- fonction(r) { return(4*pi*r^2) }
vol.cube <- fonction(r) { return(r^3) }
surf.cube <- fonction(r) { return(6*r^2) }
r <- 10^seq(1,2,le=30)
plot(log10(vol.sphere(r)),log10(surf.sphere(r)), pch=19, col="blue",
      xlab="Log10 Volume",
      ylab="Log10 Surface",
      main = "Allometrie du cube et de la sphere")
points(log10(vol.cube(r)),log10(surf.cube(r)), pch=19, col="red")
lm4 <- lm(log10(surf.sphere(r))~log10(vol.sphere(r)))
lm5 <- lm(log10(surf.cube(r))~log10(vol.cube(r)))
abline(coef = lm4$coef, col="blue")
abline(coef = lm5$coef, col="red")
text(16,5,paste("sphere y = ",round(lm4$coef[2],3)," x + ",round(lm4$coef[1],3)))
text(12,15,paste("cube y = ",round(lm5$coef[2],3)," x + ",round(lm5$coef[1],3)))
```



En effet, dans le cas d'un cube on a :

$$\text{Log}(V(r)) = \text{Log}(r^3) = 3\text{Log}(r)$$

$$\text{Log}(S(r)) = \text{Log}(6r^2) = \text{Log}(6) + 2\text{Log}(r)$$

$$\text{Log}(S(r)) = \frac{2}{3}\text{Log}(V(r)) + \text{Log}(6) = 0.667\text{Log}(V(r)) + 0.778$$

Et dans le cas d'une sphère :

$$\text{Log}(V(r)) = \text{Log}\left(\frac{4}{3}\pi r^3\right) = -\text{Log}(3) + \text{Log}(4\pi) + 3\text{Log}(r)$$

$$\text{Log}(S(r)) = \text{Log}(4\pi r^2) = \text{Log}(4\pi) + 2\text{Log}(r)$$

$$\text{Log}(S(r)) = \frac{2}{3}\text{Log}(V(r)) + \frac{2}{3}\text{Log}(3) + \frac{1}{3}\text{Log}(4\pi) = 0.667\text{Log}(V(r)) + 0.684$$

L'ordonnée à l'origine plus petite dans le cas de la sphère vient de ce que la sphère minimise la surface nécessaire pour enclore un volume donné, mais dans les deux cas la surface augmente comme une puissance $2/3$ du volume. Si la taille relative du cerveau est liée à la surface corporelle (logique, notre peau est une interface neurologique après tout) alors la taille du cerveau devrait être une fonction puissance $2/3$ de la masse corporelle. Si le facteur limitant est le métabolisme basal (logique aussi, métabolisme élevé, plus d'énergie pour maintenir le système, et plus d'énergie disponible pour des tâches complexes, comme la recherche d'une source d'énergie!) alors la fonction puissance devrait être voisine de 0.75. D'où l'intérêt de déterminer la valeur de la pente : nous indiquer ce qui est important d'un point de vue biologique.

Revenons maintenant à l'application de la méthode au problème de l'intelligence et des des cerveaux :

3.1 Que nous dit la valeur de la pente ?

Une des premières, et des plus citées, étude d'allométrie est le travail de Jerison sur la taille relative des cerveaux des vertébrés [6]. La figure ci-dessous est extraite de [5].

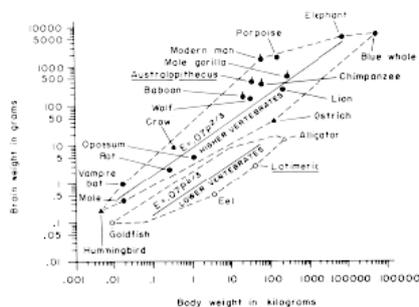
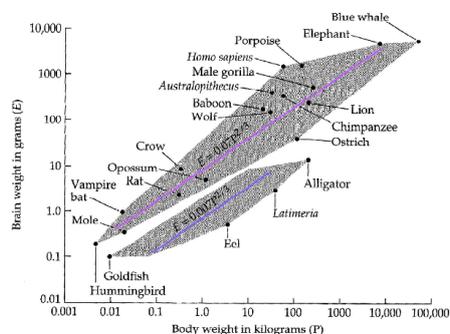


FIG. 7. Convex polygons and allometric lines at "slope" $\frac{2}{3}$ for the discriminable groups of Fig. 6 (redrawn from Jerison, 1973).

C'est une figure qui a été reprise de nombreuses fois, par exemple, avec une meilleure résolution et en couleur, celle extraite de Schlichting et Pigliucci [7] :



Les données originelles ne sont malheureusement pas disponibles sur le Web. Le décalage entre les deux droites qui de l'ordre d'une unité en Log décimal, soit un facteur 10, est censé être le reflet d'une différence fondamentale entre les vertébrés « inférieurs » et « supérieurs » en termes d'allocation de masse cerveau/corps. Cependant au sein de chaque groupe les pentes sont égales (exposant allométrique = $\frac{2}{3}$, ou 0.667) suggérant qu'à l'intérieur de chaque groupe les mêmes contraintes jouent pour la taille relative du cerveau et du corps. Comme 0.67 est l'exposant pour la relation entre la surface et le volume des solides, cela suggère que la taille relative des cerveaux des vertébrés est « pilotée » par la surface corporelle, avec les vertébrés « supérieurs » décalés vers le haut en reflet de leur plus grand investissement en intelligence.

Super, sauf que Jerison a triché. Enfin, biaisé. Comme il l'explique dans ses articles, les droites ne sont pas des droites de régression (orthogonale ou quoi que ce soit), ce sont des droites de pente 0.67 qu'il a ajusté à l'oeil aux données. Ce n'est pas une démarche très scientifique (pour sa défense c'est le premier à avoir pensé à ce genre de problème, un pionnier, et les pionniers ont souvent une première image juste sans s'encombrer des détails, d'ailleurs dans les légendes de ses figures le mot « pente » est entre guillemets). Les analyses ultérieures sur de nombreux mammifères ont donné des pentes plus proches de 0.75, en faveur d'un lien avec le métabolisme basal pour la taille du cerveau. Dans notre jeu de données avec 62 mammifères, la pente de la droite de régression orthogonale est de 0.78 :

P
[1] 0.7755009

Cependant, il y a beaucoup de différences d'un groupe taxonomique à un

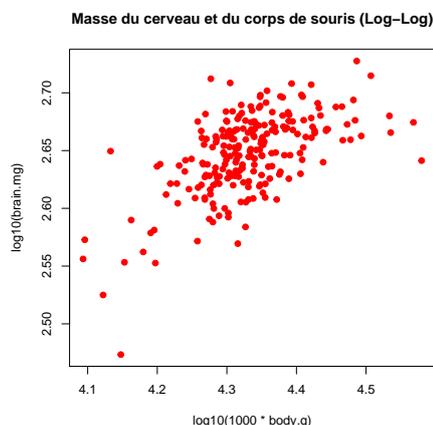
autre, et il y a de nombreux articles et des livres entiers consacrés aux problèmes de l'allométrie, la signification des pentes allométriques des masses cerveau/corps est toujours en discussion. En général :

- Pour des échantillons à large spectre taxonomique, les pentes sont voisines de 0.75, suggérant un lien avec le métabolisme.
- Quand on restreint la gamme taxonomique (mettons des mammifères, aux familles, aux espèces) les pentes tendent à décroître, jusqu'à presque s'annuler au sein des espèces : la masse corporelle explique peu les variations de taille du cerveau au sein d'une espèce.



Par exemple, le fichier `mice230.txt` comporte des données sur la masse du cerveau et la masse corporelle issues de la pesée de 4140 souris regroupées en 230 souches (source : the mouse brain library <http://www.nervenet.org/main/databases.html>, crédit photo : <http://www.biochem.usyd.edu.au/images/mice.jpg>).

```
mice230 <- read.table("http://pbil.univ-lyon1.fr/R/donnees/mice230.txt", sep='\t', header = TRUE)
attach(mice230)
plot( x = log10(1000*body.g), y = log10(brain.mg), pch = 19, col = "red",
main = "Masse du cerveau et du corps de souris (Log-Log)")
```

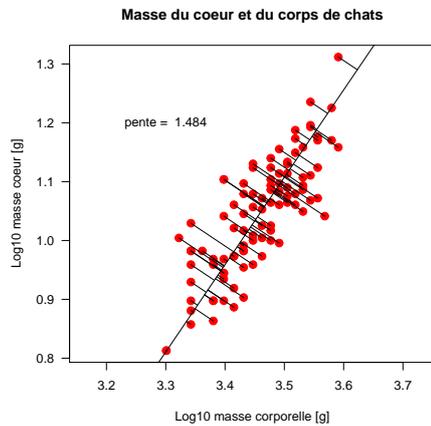


Comme on peut le constater, chez les souris, la corrélation entre la masse corporelle et la masse du cerveau est faible, comparée à la corrélation interspécifique chez les mammifères.

Attention, cela ne signifie pas qu'il n'y a pas de relations allométriques intraspécifiques pour d'autres caractères morphologiques. Par exemple, le jeu de données `catsM` de la bibliothèque `boot` donne pour 97 chats mâles adultes la masse corporelle (en kilogramme) et la masse du cœur (en gramme, et non en kilogramme comme indiqué dans la documentation de R, la masse du cœur d'un chat adulte normal ne dépassant pas les 20 grammes, voir par exemple <http://>

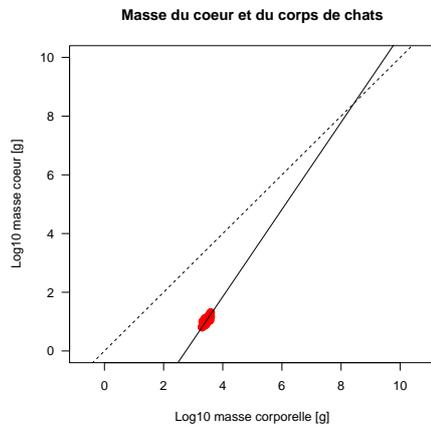
([//www.vmath.ucdavis.edu/cardio/cases/case6/followup.htm](http://www.vmath.ucdavis.edu/cardio/cases/case6/followup.htm)). Les données sont extraites de [2].

```
library(boot)
data(catsM)
attach(catsM)
x <- log10(1000*Bwt)
y <- log10(Hwt)
w <- as.data.frame(cbind(x, y))
plot(w, asp = 1, pch = 20, cex = 2, col = "red", las = 1,
     xlab="Log10 masse corporelle [g]",
     ylab="Log10 masse coeur [g]",
     main="Masse du coeur et du corps de chats")
cov <- var(w)
u <- eigen(cov, sym=TRUE)$vectors[,1]
p <- u[2]/u[1]
scal <- (x - mean(x)) * u[1] + (y - mean(y)) * u[2]
segments(x, y, mean(x) + scal * u[1], mean(y) + scal * u[2])
abline(c(mean(y) - p * mean(x), p))
text(3.3, 1.2, paste("pente = ", round(p, 3)))
```

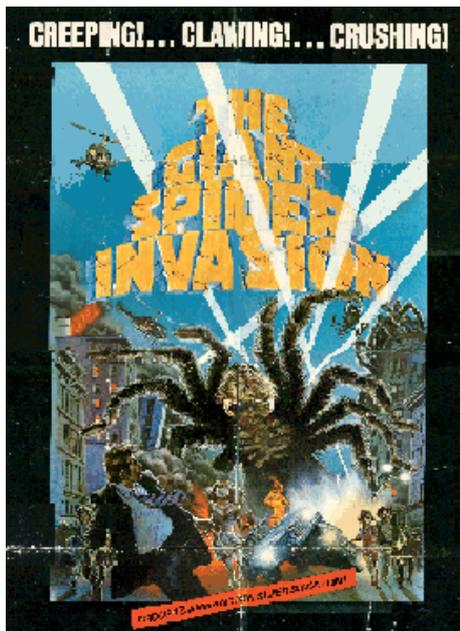


La pente étant supérieure à 1, on parle d'allométrie positive : la masse relative du coeur augmente avec la masse corporelle. En extrapolant outrageusement, nous constatons qu'un chat de 10^8 grammes, soit 100 tonnes, ne serait pas fonctionnel puisque toute la masse corporelle devait être consacrée au coeur.

```
plot(w, asp=1, pch=20, cex=2, col="red", las = 1,
     xlab="Log10 masse corporelle [g]",
     ylab="Log10 masse coeur [g]",
     xlim=c(0,10), ylim=c(0,10),
     main="Masse du coeur et du corps de chats")
abline(c(mean(y) - p * mean(x), p))
abline(coef=c(0,1), lty=2)
```



On estime que la masse maximale pour un animal terrestre (*i.e.* absence de la force d'Archimède pour contrecarrer les effets de la pesanteur) fait de chair et d'os est de l'ordre de 100 à 1000 tonnes [4].



Le non respect des proportions allométriques fait que de nombreuses formes extraterrestres inventées par simple dilatation isométrique d'espèces existantes sur terre ne sont tout simplement pas viables. Godzilla, cette sorte d'iguane dilaté jusqu'à la taille d'un gratte-ciel de Tokyo, est impossible, ses os se briseraient sous son propre poids^a. Ci-contre, l'invasion des araignées géantes de Richard Matheson (crédit photo : <http://www.mst3ktemple.com/images/posters/810-giantspider.gif>).

^a. Galilei, G. : Dialogues Concerning Two New Sciences (1637) translated by H. Crew and A. De Salvio. New York, Macmillan, 1914, 300 pp.

Conclusion : Oubliez les petits bonshommes aux pieds recouverts de fourrure et les combats à l'épée sur des planètes imaginaires⁴.

3.2 Que nous disent les écarts à la droite ?

La relation générale entre la masse corporelle et celle du cerveau, donnée par la pente (coefficient d'allométrie) peut nous apprendre des choses sur le fondement du déterminisme de la taille du cerveau pour un large spectre taxo-

⁴. *Forget the little people with fuzzy turned-up feet and sword fights on imaginary planets* Philip K. Dick (1979) Postface à Dr. Adder de K.W. Jeter ISBN 0-586-07076-1, et pour la traduction française de Michel Lederer ISBN 2-207-30409-4 (Denoël, présence du futur)

nomique. Cependant, nous sommes également intéressés par les espèces (ou les genres, etc) : *e.g.* les gorilles ont-ils un gros cerveau pour leur taille ? Sous-jacente il y a évidemment la question « sont ils intelligents » ? Il y a de nombreuses problématiques qui utilisent la taille du cerveau comme indicateur de l'intelligence et ici nous nous intéressons uniquement à la méthode, et non à savoir si elle est bien justifiée. La première intuition est que la taille absolue n'est pas un bon indicateur, sans quoi les baleines seraient les animaux les plus intelligents sur terre. Donc, nous voulons neutraliser la portion de la taille du cerveau qui est « due » à la taille du corps.

Puisque la droite de régression est calculée à partir des observations, il devrait être possible de calculer à quelle distance chaque observation se trouve de la droite, et ces distances sont appelées les résidus. On peut ensuite utiliser ces résidus d'une manière ou d'une autre pour produire un indice de la taille du cerveau par rapport au corps. Selon les chercheurs, différents indices peuvent être utilisés (les plus courants sont le *comparative brain size* ou CBS, et le *encephalization quotient* ou EQ). Les animaux avec de grands EQ sont censés être intelligents et ceux avec des petits EQ censés être ... censés ne pas l'être.

Attention : il y a de nombreuses façons de tracer une droite de régression, et comme tout cela est empirique la droite utilisée, et donc les résidus, vont dépendre des espèces qui sont dans l'échantillon, **et** il y a différentes manières de calculer un EQ (pensez-vous que le EQ humain devrait être calculé sur tous les mammifères, ou juste sur les primates ?). C'est pourquoi les valeurs publiées dans la littérature pour les EQ et les CBS varient grandement d'un auteur à l'autre. Il faut donc être très prudent avant de comparer les valeurs entre différents articles, et lire la section « matériels et méthodes » avec la plus grande attention !

4 Quelques pièges de l'allométrie

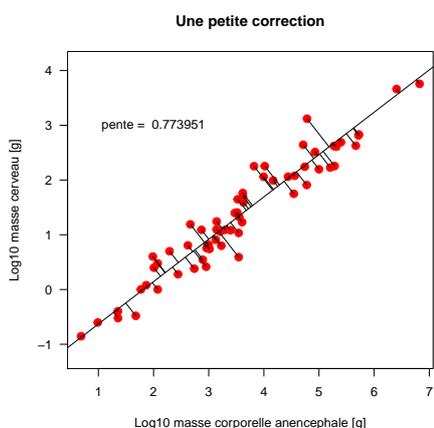
4.1 Le problème des points expérimentaux

Une *espèce* en tant que telle n'a pas de corps ni de cerveau. Combien d'individus ont-ils été mesurés ? Des animaux de zoo (souvent obèses) ? Pas de dimorphisme sexuel ? Y a-t-il une variabilité entre les populations ? Il n'y a pas de réponse simple. Lisez les « matériels et méthodes ».

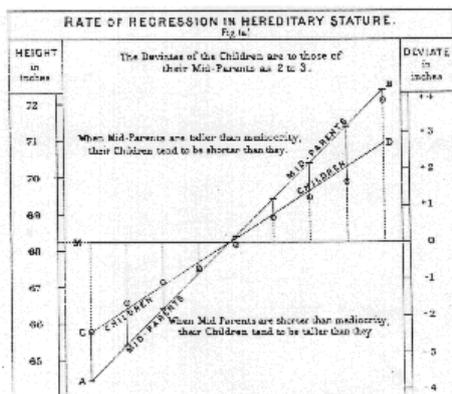
4.2 Le problème du tout et de la partie

A proprement travailler, il n'est pas légitime de corrélérer la masse du cerveau avec la masse corporelle totale parce la masse corporelle inclut la masse du cerveau, ce qui risque d'augmenter artificiellement la corrélation entre les « deux » variables. Le cerveau des primates ne représente pas une fraction si importante du corps pour que cela biaise les résultats, mais si vous vous intéressez par exemple à la masse musculaire, clairement il faudra comparer la masse musculaire non à la masse corporelle, mais à la masse corporelle hors muscles (recherche bien crade ça). Dans le cas de notre jeu de données sur 62 mammifères, cela ne change que le troisième chiffre significatif de la valeur de la pente, c'est parfaitement négligeable :

```
x <- log10(1000*body - brain)
y <- log10(brain)
w <- as.data.frame(cbind(x,y))
plot(w,asp=1,pch=20, cex=2, col="red", las = 1,
      xlab="Log10 masse corporelle anencephale [g]",
      ylab="Log10 masse cerveau [g]",
      main="Une petite correction")
cov <- var(w)
u <- eigen(cov, sym=TRUE)$vectors[,1]
p <- u[2]/u[1]
scal <- (x - mean(x) ) * u[1] + (y - mean(y) ) * u[2]
segments(x, y, mean(x) + scal*u[1], mean(y) +scal*u[2])
abline( c(mean(y) -p*mean(x), p))
text(2, 3, paste("pente = ", round(p,6)))
```



4.3 Le problème de Galton



Francis Galton est un des premiers à avoir utilisé des régressions en biologie. La figure ci-contre est extraite de [3]. Galton s'intéressait à la relation entre la taille des enfants et celle de leur parents. Constatant que la taille des enfants avait tendance à se rapprocher de la moyenne, il qualifia le phénomène du très discutable « régression vers la médiocrité », c'est incidemment de là que nous vient le terme de « droite de régression ». Crédit photo : <http://www.stat.ucla.edu/history/>.

Le problème de Galton, auquel nous sommes toujours confronté, est que les observations ne sont pas indépendantes : toutes les formes de vie actuelles dérivent d'un même ancêtre commun. Dès lors, la ressemblance entre deux individus peut n'être que le reflet artefactuel d'une simple contingence historique, la mémoire d'une condition initiale commune. Il existe des méthodes statistiques pour pallier ce type de problème à la condition express de disposer d'un arbre décrivant les relations phylogénétiques entre les individus.

4.4 Le problème de l'interprétation

Un exemple simple : si vous classez les primates en fonction de leur régime alimentaire en distinguant les herbivores (mangeurs de plantes) des frugivores (mangeurs de fruits), et que vous faites la régression de la taille du cerveau sur celle du corps, les herbivores ont tendance à être sous la droite de régression, ils ont tendance à avoir un EQ plus faible que les frugivores. La taille relativement petite des cerveaux des herbivores a été utilisée comme argument en faveur de l'existence d'une pression écologique forte pour l'évolution de l'intelligence. **Mais** : cela suppose que la taille relative du cerveau soit une mesure de l'intelligence (quand on compare des **espèces différentes** ce n'est pas complètement absurde). Cela suppose également que rien d'autre n'entre en jeu ; et plusieurs difficultés liés à cette hypothèse ont été soulevées :

- les herbivores peuvent avoir un métabolisme basal plus faible, responsable de la taille de leur cerveau, sans qu'il n'y ait un manque de besoin pour l'intelligence ;
- les herbivores ont de gros systèmes digestifs, pour assurer la fermentation des végétaux, leur masse est peut-être biaisée par cette augmentation « artificielle » de leur taille, diminuant la masse relative de leur cerveau.

Comment comparer de gras herbivores avec de maigres frugivores ?

Références

- [1] T. Allison and D.V. Cicchetti. Sleep in mammals : ecological and constitutional correlates. *Science*, 194 :732–734, 1976.
- [2] R.A. Fisher. The analysis of covariance method for the relation between a part and the whole. *Biometrics*, 3 :65–68, 1947.
- [3] F. Galton. Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute*, 15 :246–263, 1886.
- [4] J.E.I. Hokkanen. The size of the largest land animal. *Journal of Theoretical Biology*, 118 :491–499, 1986.
- [5] H.F. Jerison. The evolution of the mammalian brain as an information-processing system. In J.F. Eisenberg and D.G. Kleiman, editors, *Advances in the Study of Mammalian Behavior (Spec. Publ. Amer. Soc. Mamm. 7)*, pages 113–146, Pittsburgh, USA, 1983. American Society of Mammalogists.
- [6] H.J. Jerison. *Evolution of the Brain and Intelligence*. Academic Press, New York, USA, 1973. Plus de 680 citations au 31-OCT-2004.
- [7] C.D. Schlichting and M. M. Pigliucci. *Phenotypic Evolution : A Reaction Norm Perspective*. Sinauer Associates, Sunderland, MA., USA, 1998.
- [8] David Warton and translated to R by John Ormerod. *smatr : (Standardised) Major Axis Estimation and Testing Routines*, 2007. R package version 2.1.
- [9] D.I. Warton, I.J. Wright, D.S. Falster, and M. Westoby. Bivariate line-fitting methods for allometry. *Biol. Rev.*, 81 :259–291, 2006.