

## Relations dose-effet

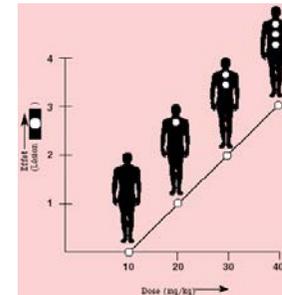
Marc Girondot, Université Paris Saclay  
marc.girondot@universite-paris-saclay.fr

Version du 03/12/2020

1

## Principe en toxicologie

- Un principe important en toxicologie veut que toutes les substances chimiques soient toxiques, car il existe toujours une dose pouvant causer un effet nocif. Mais le fait d'inhaler, de toucher et même d'ingérer des substances chimiques n'entraîne pas nécessairement l'apparition d'un tel effet.



2

## Relation dose-effet

- La notion de seuil toxique est importante, car elle peut servir à fixer des normes.
- La valeur seuil représente la quantité minimale sous laquelle il ne se produit pas d'effet. Au-dessus de ce seuil, l'effet observé dépend de la dose.

3

Paracelse, né Philippus Theophrastus Aureolus Bombast von Hohenheim en 1493 à Einsiedeln (en Suisse centrale) et mort le 24 septembre 1541 à Salzbourg



- Avancées: Analyse réductionniste des maladies, extraction des principes actifs des substances, usage interne des médicaments chimiques.
- La pensée de Paracelse est le point de départ du long processus de séparation de la chimie de l'alchimie.
  - L'alchimie est une discipline qui peut se définir comme « un ensemble de pratiques et de spéculations en rapport avec la transmutation des métaux ». Cet objectif se fonde sur la théorie que les métaux sont des corps composés (souvent de soufre et de mercure).
  - Un autre objectif classique de l'alchimie est la recherche de la panacée (médecine universelle) et la prolongation de la vie via un élixir de longue vie.

4

Vénus

## La dose et le poison



- La syphilis (ou vérole), une maladie nouvelle qui avait explosé en Europe à la fin du XV<sup>ème</sup> siècle, était traitée par les charlatans, barbiers et chirurgiens, avec le mercure. Paracelse réagit d'abord en déclarant qu'elle était due à la licence sexuelle des hommes et à une configuration astrale de Vénus qui transmutait les anciennes maladies en nouvelles maladies.
- Pour atténuer les symptômes cutanés, il utilisait un onguent mercuriel, remède connu dès le Moyen-Âge. Paracelse a vu que le mercure mal dosé, tue.
- La citation célèbre, du médecin (al)chimiste est:

**Toutes les choses sont poison, et rien n'est sans poison ; seule la dose fait qu'une chose n'est pas poison.**

Aphrodite, dite Venus de Milo

5

## La syphilis

- La syphilis (connue familièrement sous le nom de vérole ou encore de grande vérole par opposition à la variole) est une infection sexuellement transmissible contagieuse, due à une bactérie.
- Une étude portant sur des ossements médiévaux aurait établi qu'une forme bénigne de la syphilis existait alors en Europe, mais l'apparition d'une forme mutante virulente originaire probablement du Nouveau Monde apparaît en Europe peu de temps après le retour des marins de Christophe Colomb.



Kaare Lund Rasmussen, Jesper Lier Boldsen, Hans Krøgaard Kristensen, Lilian Skytte, Katrine Lykke Hansen, Louise Molholm, Pieter M. Grootes, Marie-Josée Nadeau, Karen Marie Fløche Eriksen, « Mercury levels in Danish Medieval human bones », Journal of Archaeological Science, vol. 35, no 8, août 2006, p. 2295-2306

6

## La syphilis

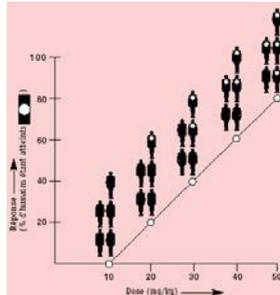
- La maladie a été décrite dès la fin du XV<sup>ème</sup> siècle mais le tréponème pâle a été identifié par Fritz Schaudinn et Erich Hoffmann à Berlin en 1905.
- Le mercure fut le remède pluricentenaire, même si son efficacité n'a jamais été démontrée. Le traitement réel apparaît au début des années 1940, avec l'avènement des antibiotiques, en particulier l'usage de la pénicilline.



7

## Application à une population

- Le même principe s'applique à une population d'individus, car l'effet ou les nombreux effets possibles peuvent se manifester différemment chez plusieurs personnes exposées à une même dose d'un toxique. C'est ce qu'on appelle la relation dose-réponse ou exposition-réponse, soit la relation entre l'exposition et le nombre d'individus qui présentent un effet donné.



8

## Synthèse

- Une augmentation de la dose peut entraîner une augmentation des effets chez un individu ;
- Une augmentation de la dose peut augmenter la proportion des individus affectés.

9

# RELATION DOSE-EFFET

10

## Les enjeux

- La relation dose-effet ou relation exposition-réponse ou plus simplement écrite dose-réponse exprime le changement d'effet, sur un organisme, provoqué par une quantité différente de « stresseurs », ou de « stimuli » après un certain temps d'exposition.
- Cette notion est l'une des bases de l'établissement de « niveaux » et « seuils d'intervention » réglementaires face aux contaminants jugés les plus préoccupants, néanmoins pondérée par d'autres éléments tels que les conditions techniques et économiques du moment.

11

## Dose reçue

- La **biodisponibilité** est la proportion d'une substance qui va effectivement agir dans le système par rapport à la quantité présente.
  - Propriétés des substances ;
  - Propriétés du compartiment de l'environnement considéré (ex : l'acidité augmente la biodisponibilité des métaux, de même que le potentiel redox. La granulométrie du sol et la turbidité de l'eau font que les particules vont ou non plus ou moins adsorber certains contaminants, et d'autres facteurs biotique/abiotique interviennent (ex : selon que la nature du substrat pollué favorise ou non la bioturbation)...

12

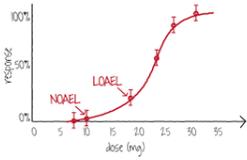
### Courbe dose-réponse

- On cherche pour chaque contaminant à établir ou modéliser une relation dose-réponse, souvent sous forme de « courbe » dose-réponse.
- Une telle courbe peut traduire une relation « linéaire sans seuil », ou au contraire comprendre des ruptures (effet de seuil, toxicité aiguë, mort).

13

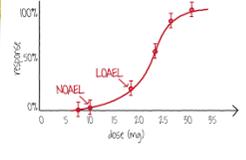
### Courbe avec seuil: NOAEL

- La dose sans effet nocif observable (DSENO), appelée aussi « dose sans effet toxique », « dose maximale sans effet » ou « dose maximale sans effet néfaste observable » (en anglais : No Observable Adverse Effect Level, NOAEL), est la dose la plus élevée d'une substance chimique ne produisant aucun effet nocif observable au cours d'une étude de toxicité.



14

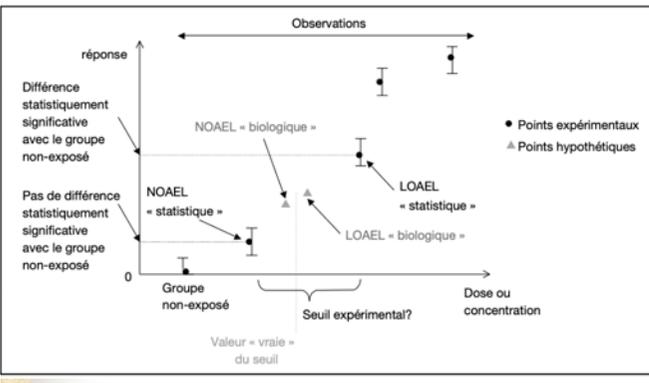
### Courbe avec seuil: LOAEL



- L'étude expérimentale ne permet pas toujours d'avoir accès à ce NOAEL. Il est alors proposé de déterminer la dose ou la concentration théoriquement la plus faible pour laquelle un effet indésirable est observé. C'est la dose minimale pour un effet nocif observable (DMENO ou LOAEL en anglais pour Lowest Observed Adverse Effect Level)
- Plus précisément, elle correspond à la plus faible dose de substance pour laquelle on constate une augmentation statistiquement (ou biologiquement) significative en fréquence ou en sévérité d'un effet nocif observé dans le groupe exposé par rapport au groupe non exposé.

15

### NOAEL et LOAEL



16

## Benchmark dose (ou BMD)

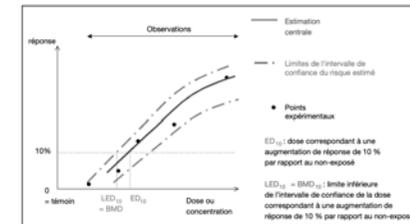
- L'objectif est de déterminer la dose (ou la limite inférieure de l'intervalle de confiance correspondant) produisant un effet critique avec une augmentation de la fréquence ou de la sévérité particulière, conventionnellement fixée à 1, 5 ou 10 %. Le principe de cette méthode décrite par Crump en 1984 repose sur un ajustement statistique de la totalité des données d'observation.
- Ce n'est pas une méthode permettant l'extrapolation aux faibles doses mais seulement une méthode d'estimation de la dose critique.

Crump KW, A new method for determining allowable daily intake, *Fundamental Applied Toxicology*, 1984, 4 : 854-871.

17

## Benchmark dose (ou BMD)

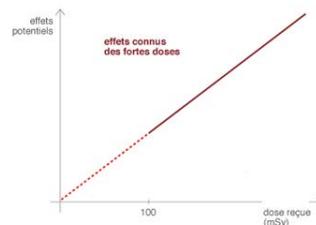
- Par définition, la réponse fixée pour déterminer une BMD est la réponse la plus basse pouvant être détectée expérimentalement, mais suffisante pour que la BMD soit relativement insensible au choix du modèle statistique. Ce compromis permet d'utiliser les observations pour les niveaux de doses les plus faibles possibles.



18

## Courbe linéaire sans seuil

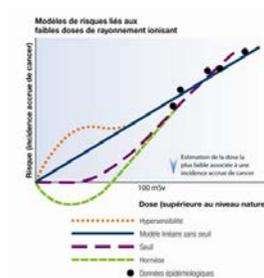
- Ce modèle est notamment appliqué dans la radioprotection. Ce modèle suppose qu'il existe un lien direct et proportionnel entre l'exposition au rayonnement et les risques de cancer pour toutes les doses de rayonnement.



19

## Extrapolation

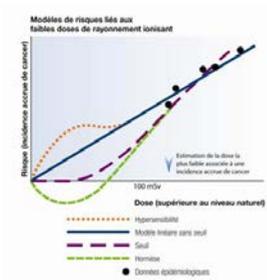
- En dehors de la zone avec des observations, le comportement de la courbe est mal connu.
  - Le modèle basé sur l'hypersensibilité suggère un risque plus important à doses plus faibles.
  - Le modèle LSS est la ligne droite extrapolée jusqu'à zéro, c.-à-d. que le risque de cancer s'accroît avec la hausse de la dose.
  - Le modèle fondé sur un seuil sous-entend que sous une certaine dose, il n'existe pas de risque.
  - Le modèle fondé sur l'hormèse suggère que de faibles doses de rayonnement peuvent même avoir un effet protecteur et bénéfique.



20

## Absence de seuil ou absence de connaissance ?

- La partie de la courbe à gauche correspond à une zone où la qualité des estimations de la variable indépendante (le polluant) est la plus faible et celle où la variable dépendante (la réponse) est la plus faible, donc la moins connue !



21

## Les recherches actuelles : L'hormèse

Friedrich Nietzsche : « À l'école de guerre de la vie. — Ce qui ne me fait pas mourir me rend plus fort. » dans *Le Crépuscule des idoles* (1888)

- En toxicologie, le phénomène d'hormèse se caractérise par une forme caractéristique de la courbe de relation dose / effet, qui change de signe pour les faibles doses, ce qui lui donne une forme en « U » ou en « J ».
- L'idée générale que de faibles doses puissent avoir des effets différents des doses fortes (et parfois radicalement différent) est connue et acceptée, mais cela ne signifie pas nécessairement que l'effet de la faible dose soit toujours bénéfique.
  - Il y a une ambiguïté sur ce dernier point. Le terme hormèse doit-il s'appliquer seulement au cas où les faibles doses protègent ou bien au cas plus généraux où la réponse n'est pas monotone ? Selon les auteurs, les deux acceptions sont notées.

22

## Les recherches actuelles : L'hormèse

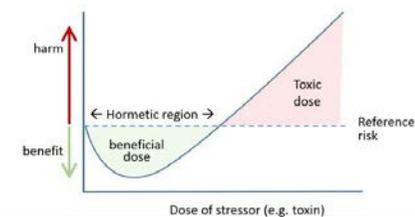
- Les mécanismes par lesquels l'hormèse (au sens réponse bénéfique) se manifeste ne sont pas bien compris. On pense globalement que la présence d'une faible dose de toxique déclenche certains mécanismes d'auto-réparation dans la cellule ou l'organisme, et ces mécanismes une fois activés sont suffisants pour non seulement neutraliser l'effet initial du toxique, mais également réparer d'autres défauts que le toxique n'avait pas provoqués.

Vandenberg, L. N., et al. (2012). "Hormones and endocrine-disrupting chemicals: low-dose effects and nonmonotonic dose responses." *Endocr Rev* 33(3): 378-455.

23

## Conséquences réglementaires

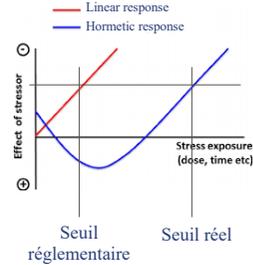
- Si l'hormèse est vraiment relativement courante, un enjeu de santé publique et de santé environnementale est de l'intégrer aux pratiques d'évaluation toxicologiques, puis à la réglementation des toxiques et aux pratiques d'évaluation et de gestion des risques.



24

## Conséquences réglementaires

- L'approche « linéaire sans seuil » est une approche majorante, dictée par le principe de précaution faute d'une meilleure information : même si l'on a des raisons de penser que d'autres phénomènes peuvent apparaître aux faibles doses, il n'y a pas lieu de retenir un autre modèle tant que l'existence d'un seuil n'est pas clairement établie.
- Cependant, lorsque les études montrent une relation dose / effet non linéaire, des modèles à seuil (impliquant par exemple une absence de risque de cancer à des doses inférieures à un seuil) sont couramment acceptés.



25

## Cas particulier des perturbateurs endocriniens

- Les perturbateurs endocriniens sont des substances qui dérèglent le fonctionnement hormonal des organismes vivants et peuvent entraîner ainsi des effets néfastes sur la santé et l'environnement.
- "Un perturbateur endocrinien est une substance ou un mélange de substances, qui altère les fonctions du système endocrinien et de ce fait induit des effets néfastes dans un organisme intact, chez sa progéniture ou au sein de (sous)- populations". OMS 2002
- Les perturbateurs endocriniens peuvent interférer avec toutes les grandes fonctions des organismes vivants : croissance, reproduction, comportement, nutrition, métabolisme, système nerveux...

26

## Faibles doses d'exposition

- Habituellement, en dessous d'un certain niveau d'exposition, les mécanismes de défense de l'organisme permettent d'éviter l'apparition d'effets sanitaires. On parle alors d'effet de seuil.
- Pour certaines substances dangereuses comme des molécules cancérigènes, on observe qu'il n'y a parfois pas d'effet de seuil, au moins à l'échelle d'une population donc, des effets possibles même à faible dose.
- Les perturbateurs endocriniens sont suspectés d'agir de même.

27

## Les relations dose-réponse non monotones

- Traditionnellement, les effets nocifs des substances chimiques sont décrits, dans les études de toxicologie comme proportionnel à la dose testée. Typiquement, une faible dose ne produit pas d'effet, la dose médiane produit de faibles effets toxiques alors que la forte dose testée induit des effets plus prononcés ou plus délétères.
- Mais, certaines substances chimiques peuvent suivre des courbes inversées, c'est-à-dire avoir des effets plus importants à faible dose à ceux observés à fortes doses, on parle alors de dose réponse non-monotone.

28

## Modéliser l'effet

- Comme signalé précédemment, les dosages des faibles doses présentent une marge d'erreur importante et par ailleurs l'étude des effets sont aussi plus difficiles à mettre en œuvre car ils sont rares.
- Il existe donc une hétéroscédasticité importante...

L'hétéroscédasticité correspond à une non-indépendance entre la moyenne et la variance. Cela s'oppose à l'homoscédasticité.

29

## DÉCRIRE LES DONNÉES

30

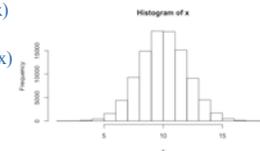
## Le nombre magique 2

- On a tous entendu que 95% des points sont situés dans l'intervalle moyenne  $\pm 2$  SD.
- Déjà, on sait tous que ce n'est pas 2, mais 1,96:  
> (deux <- qnorm(p=0.975, 0, 1))  
[1] 1.959964
- Si on prend 2, cela fait:  
> pnorm(2, 0, 1)-pnorm(-2, 0, 1)  
[1] 0.9544997
- Pas si différent si l'effectif n'est pas trop important.

31

## Distribution des observations

```
> x <- rnorm(100000, 10, 2)
> hist(x)
> mean(x)
[1] 10.01199
> sqrt(sum((x-mean(x))^2)/length(x))
[1] 1.997326
> sd(x)
[1] 1.997336
> sum(x < (mean(x)-deux*sd(x)))/length(x)
[1] 0.02401
> sum(x > (mean(x)+deux*sd(x)))/length(x)
[1] 0.02531
```



32

### Distribution des observations

```

> x <- rlnorm(100000, 6, 0.5)
> hist(x)
> mean(x)
[1] 457.4219
> sd(x)
[1] 245.5845
> sum(x<(mean(x)-deux*sd(x)))/length(x)
[1] 0
> sum(x>(mean(x)+deux*sd(x)))/length(x)
[1] 0.04643

```

33

### Distribution des observations

```

> x <- runif(100000, 5, 15)
> hist(x)
> mean(x)
[1] 9.994812
> sd(x)
[1] 2.890294
> sum(x<(mean(x)-deux*sd(x)))/length(x)
[1] 0
> sum(x>(mean(x)+deux*sd(x)))/length(x)
[1] 0

```

34

### Distribution des observations

```

> x <- rbinom(100000, 10, 0.1)
> hist(x)
> mean(x)
[1] 1.00094
> sd(x)
[1] 0.9483818
> sum(x<(mean(x)-deux*sd(x)))/length(x)
[1] 0
> sum(x>(mean(x)+deux*sd(x)))/length(x)
[1] 0.07043

```

35

### Conclusion

- 95% des observations ne sont pas dans l'intervalle  $\pm 1.96$  SD sauf dans un cas très particulier: la loi normale.
- Cela peut conduire à des situations à la limite de l'absurde:

```

x <- x/10
mean(x)
sd(x)
library(HelpersMG)

plot_errbar(1, mean(x), errbar.y = deux*sd(x), bty="n",
            ylim=c(-0.1, 1), xlab="", ylab="Fréquence", xaxt="n")
segments(x0=0, x1=2, y0=0, y1=0, lty=2)

```

36

## Probabilité, fréquence, vraisemblance

- Il existe de nombreuses définitions de ces termes qui ne recouvrent donc pas exactement la même chose chez différents auteurs.
- En général, le terme probabilité est utilisé pour décrire la chance (ou le risque !) qu'un événement particulier se produise.
- Le terme fréquence (ou plus précisément fréquence empirique) désigne une réalisation particulière de la probabilité qu'un événement particulier se produise.
- Le terme vraisemblance désigne une probabilité conditionnelle, c'est à dire la probabilité d'observations en fonction de paramètres supposés connus.
- En anglais, *frequency* est utilisé pour nommer des dénombrements (par exemple sur la commande `hist()` de R, `freq=TRUE` permet d'afficher les nombres d'observations); on distingue:
  - *Frequency (or absolut frequency)*, des nombres
  - *Relative frequency (or empirical probability)*, des fréquences !

37

## Masquer les erreurs

- Il existe une infinité de mesures de la dispersion, les plus fréquemment utilisées étant l'écart-type et l'erreur standard.
- On a comme relation :
- Donc SE est toujours plus faible que SD; et donc autant présenter SE que SD sur les barres d'erreur, cela cachera qu'on a mal travaillé !

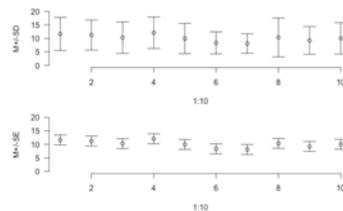
38

## Exemple SE vs SD

```
M <- NULL
SD <- NULL
SE <- NULL
```

```
for (i in 1:10) {
  x <- runif(10, 5, 15)
  M <- c(M, mean(x))
  SD <- c(SD, sd(x))
  SE <- c(sd(x)/sqrt(length(x)))
}
```

```
library(HelpersMG)
par(mar=c(4, 4, 1, 1))
layout(mat = matrix(1:2, nrow=2))
plot_errbar(1:10, M, errbar.y = deux*SD, bty="n", las=1, ylab="M+/-SD", xlab="1:10", ylim=c(0, 20))
plot_errbar(1:10, M, errbar.y = deux*SE, bty="n", las=1, ylab="M+/-SE", xlab="1:10", ylim=c(0, 20))
```



39

## Barres d'erreur

- Il y a une honte à présenter des barres d'erreur très grandes.
- En fait le terme « barre d'erreurs » serait à proscrire. Ces barres mesurent la dispersion d'un résultat, pas forcément une erreur.
- La source de dispersion peut être due à une variabilité naturelle, à une variabilité générée par la mesure ou bien réellement à des erreurs de mesure.
- Vouloir masquer la variabilité naturelle est stupide puisque justement c'est cette variabilité naturelle qui est utile pour répondre à des questions intéressantes (ou pas) !

40

## SD vs SE... qui gagne

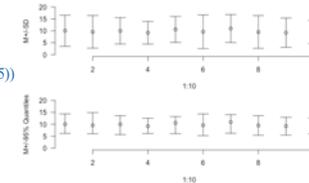
- SD est une mesure de la dispersion des observations.
- SE est une mesure de la dispersion des moyennes, c'est à dire où se trouve la moyenne
- SD et SE ne sont donc pas interchangeable; selon ce qu'on veut montrer sur un graphique, il faut choisir l'un ou l'autre.
- Il existe d'autres mesures de la dispersion qui sont souvent plus adéquates: les quantiles
- Q<sup>ème</sup> quantile: Valeur pour laquelle 1/Q observations sont en dessous de cette valeur

41

## Les quantiles 0,025 et 0,975

```
M <- NULL; SD <- NULL; SE <- NULL; QM <- NULL; QP <- NULL
```

```
for (i in 1:10) {  
  x <- runif(10, 5, 15)  
  M <- c(M, mean(x))  
  l <- quantile(x = x, probs = c(0.025, 0.975))  
  QM <- c(QM, l[1])  
  QP <- c(QP, l[2])  
  SD <- c(SD, sd(x))  
  SE <- c(SE, sd(x)/sqrt(length(x)))  
}
```

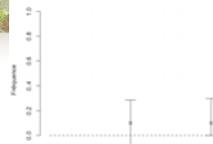


```
layout(mat = matrix(1:2, nrow=2))  
plot_errbar(1:10, M, errbar.y = deux*SD, bty="n", las=1, ylab="M+/-SD", xlab="1:10", ylim=c(0, 20))  
plot_errbar(1:10, M, y.plus = QP, y.minus = QM, bty="n", las=1, ylab="M+/-95%  
Quantiles", xlab="1:10", ylim=c(0, 20))
```

42

## Retour aux fréquences

```
x <- rbinom(100000, 10, 0.1)  
x <- x/10  
mean(x)  
sd(x)  
l <- quantile(x = x, probs = c(0.025, 0.975))  
library(HelpersMG)  
plot_errbar(c(1, 2), c(mean(x), mean(x)), errbar.y.plus = c(deux*sd(x), l[2]-  
mean(x)), errbar.y.minus = c(deux*sd(x), mean(x)-l[1]), bty="n", ylim=c(-0.1,  
1), xlab="", ylab="Fréquence", xaxt="n", xlim=c(0, 3))  
segments(x0=0, x1=2, y0=0, y1=0, lty=2)
```



43

## Distribution de la moyenne

- « La distribution de la moyenne est normale même si la distribution de la variable aléatoire ne l'est pas. »
- On retrouve cette assertion sous différentes formes notamment que les estimations obtenues par maximum de vraisemblance sont distribuées normalement.
- Qu'en est-il ?

44

### Distribution de la moyenne

```
M <- NULL; SD <- NULL; SE <- NULL

for (i in 1:10000) {
  x <- rlnorm(10, 6, 0.5)
  M <- c(M, mean(x))
  SD <- c(SD, sd(x))
  SE <- c(sd(x)/sqrt(length(x)))
}

layout(mat = matrix(1:2, nrow=2))
hist(rlnorm(100000, 6, 0.5))
hist(M, xlim=c(200, 1000))
```

```
> mean((M-mean(M))/sd(M))^3
[1] 0.5624483
> library(e1071)
> skewness(M, type=3)
[1] 0.5624483
```

45

### Distribution de la moyenne

```
M <- NULL
SD <- NULL
SE <- NULL

for (i in 1:10000) {
  x <- log(rnorm(10, 6, 1))
  M <- c(M, mean(x))
  SD <- c(SD, sd(x))
  SE <- c(sd(x)/sqrt(length(x)))
}
```

```
> skewness(M, type=3)
[1] -0.1442254
```

46

### D'où vient cette affirmation ?

- Le théorème central limite (Laplace, 1809) établit la convergence en loi de la somme d'une suite de variables aléatoires vers la loi normale.
- La moyenne est une somme divisée par une constante, donc le théorème central limite s'applique mais que dit-il en clair:
 
$$\lim_n \mathbb{E}[\varphi(X_n)] = \mathbb{E}[\varphi(X)].$$
- Définition — On dit que la suite  $(X_n)_{n \geq 1}$  converge en loi vers  $X$  si, pour toute fonction  $\varphi$  continue bornée sur  $E$ , à valeurs dans  $\mathbb{R}$ .
- Le théorème central limite nous dit donc que les estimateurs sont distribués normalement mais on ne l'atteint réellement que quand un nombre très grand d'observations utilisés pour estimer les distributions.

Pierre-Simon Laplace, « Mémoire sur les approximations des formules qui sont fonctions de très-grands nombres, et sur leur application aux probabilités », Mémoires de la Classe des sciences mathématiques et physiques de l'Institut de France, 1809, p. 353-415

47

## COMMENT MODÉLISER LES DONNÉES

48

## Les données

- Les données collectées sont en général, pour chaque individu:
  - Des données biométriques ou d'histoire de vie
  - Des données de concentration d'un ou de polluants
  - Un ou des effets:
    - Reproduction, mortalité, prise de masse
- Les effets potentiels sont les variables dépendantes, les données biométriques et les concentrations sont des variables indépendants.

49

## Maximum de vraisemblance

- Utilisation du maximum de vraisemblance pour déterminer les paramètres d'une distribution.
- Etant donné un échantillon observé  $x$ , des paramètres  $\theta$  et une fonction de vraisemblance  $f(x|\theta)$ , la vraisemblance quantifie la probabilité que les observations  $x$  proviennent effectivement d'un échantillon de la loi  $f$  paramétrée avec  $\theta$ .
- Estimer un paramètre par la méthode du maximum de vraisemblance, c'est proposer comme valeur des paramètres  $\theta$  ceux qui rendent maximale la vraisemblance, à savoir la probabilité d'observer les données  $x$  comme réalisation d'un échantillon de la loi  $f$ .
- Elle s'exprime à partir de la fonction de densité  $f(x|\theta)$  par

$$L(x_1, x_2, \dots | \theta) = \prod_{i=1}^n f(x_i; \theta)$$

50

## Estimation du modèle par maximum de vraisemblance

- Le terme vraisemblance désigne une probabilité conditionnelle, c'est à dire la probabilité d'observations en fonction de paramètres.
- Exemple:
  - Quelle est la probabilité d'observer une valeur de  $Y_i$  sachant que cette observation est issue d'un modèle ayant une valeur moyenne  $\mu$  et d'écart-type  $\sigma$  ?

51

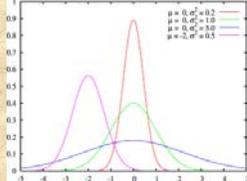
Pierre-Simon de Laplace, Essai sur la philosophie des probabilités, 1820

Loi normale ou de Gauss-Laplace

Laplace (1749-1827) et Gauss (1777-1855)



- Soit une loi normale  $\mathcal{N}(\mu, \sigma)$
- La vraisemblance d'une observation  $Y_i$  tirée dans cette loi normale est



$$L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$$

$$-\ln L\{Y; m\}$$

Pourquoi travailler en  $\ln$  ?

52

### logarithmes

$\log_a(xy) = \log_a(x) + \log_a(y)$   
 $\log_a(a) = 1 \quad \log_a(1) = 0$   
 $\log_a(x/y) = \log_a(x) - \log_a(y)$   
 $\log_a(x^y) = y \log_a(x)$

|          |           |   |
|----------|-----------|---|
| $p$      | 0         | 1 |
| $\ln p$  | $-\infty$ | 0 |
| $-\ln p$ | $+\infty$ | 0 |

« Maximum de vraisemblance »

Ces règles s'appliquent quelque soit la base  $a$  choisie pour le logarithme.

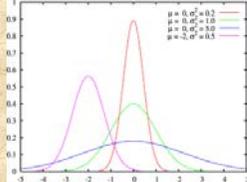
53

Pierre-Simon de Laplace, Essai sur la philosophie des probabilités, 1820

### Loi normale ou de Gauss-Laplace

Laplace (1749-1827) et Gauss (1777-1855)

- Soit une loi normale  $\mathcal{N}(\mu, \sigma)$
- La vraisemblance d'une observation  $Y_i$  tirée dans cette loi normale est



$$L\{Y_i; m\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(Y_i - \mu)^2}{2\sigma^2}\right)$$

$$-\ln L\{Y_i; m\} = \ln(\sigma) + \frac{1}{2} \ln(2\pi) + \frac{(Y_i - \mu)^2}{2\sigma^2}$$

54

Soit un échantillon de  $k$  points

Si  $\sigma$  est constant

$$-\ln L\{Y; m\} = k \left( \ln(\sigma) + \frac{1}{2} \ln(2\pi) \right) + \frac{1}{2\sigma^2} \sum_{i=1}^k (Y_i - \mu)^2$$

Notez que le premier terme est une constante qui ne dépend pas des observations.  
 Le second terme est une somme des moindres carrés pondérés par la variance qui est elle-même une constante.

Distribution  $\gamma \quad f(x; \alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)}$  for  $x > 0 \quad \alpha, \beta > 0$ ,

where  $\Gamma(\alpha)$  is the gamma function. For all positive integers,  $\Gamma(\alpha) = (\alpha - 1)!$ .

55

### Test des variables indépendantes pour expliquer la variable dépendante

- On écrit un modèle du type :

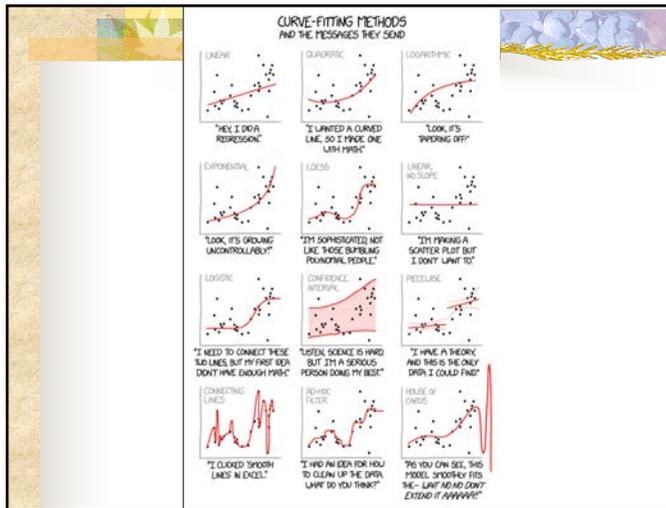
Effet  $Y = f(\text{polluant}, \text{biométrie1}, \text{biométrie2}, \text{biométrie3}, \text{etc})$

Comment choisir  $f$  ?

|   |     |
|---|-----|
| $Y = f(\text{polluant}, \text{biométrie1}, \text{biométrie2}, \text{biométrie3})$ | ML1 |
| $Y = f(\text{biométrie1}, \text{biométrie2}, \text{biométrie3})$                  | ML2 |
| $Y = f(\text{polluant}, \text{biométrie2}, \text{biométrie3})$                    | ML3 |
| $Y = f(\text{polluant}, \text{biométrie1}, \text{biométrie3})$                    | ML4 |

Etc.

56



57

### Qualité d'un ajustement

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \in [0, 1]$$

- Le coefficient de détermination ( $R^2$ ) est une mesure de la capacité du modèle à prédire les valeurs ( $y_i$  est la valeur observée,  $\hat{y}_i$  est la valeur prédite). Comment savoir qu'un ajustement est meilleur qu'un autre ?
  - $R^2$  plus fort

Mais la valeur du critère d'ajustement est dépendante du nombre de paramètres dans le modèle.

- Si on a autant de paramètres que d'observations, l'ajustement sera parfait :
  - $R^2$  vaudra 1 !

Spieß, A.N., Neumeier, N., 2010. An evaluation of  $R^2$  as an inadequate measure for nonlinear models in pharmacological and biochemical research: a Monte Carlo approach. *BMC Pharmacol* 10, 6.

58

### Le rasoir d'Ockham



- Le rasoir d'Ockham est un principe de raisonnement philosophique entrant dans les concepts de rationalisme. Son nom vient du philosophe franciscain Guillaume d'Ockham (1285, Ockham, Royaume-Uni - 1347, Munich, Allemagne).
- On le trouve également appelé principe de simplicité, principe d'économie ou principe de parcimonie (en latin *lex parsimoniae*).

« *Les multiples ne doivent pas être utilisés sans nécessité.* »

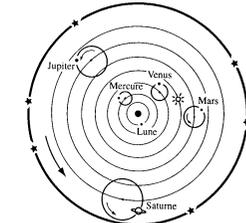
Il reprend un adage scolaire dérivé d'Aristote : « *C'est en vain que l'on fait avec plusieurs ce que l'on peut faire avec un petit nombre.* »

59

Claude Ptolémée, né vers 100 - mort vers 168 à Canope (Egypte).

### Modèle géocentrique

- Le mouvement des planètes est expliqué par Ptolémée par une série de mouvements circulaires avec la terre au centre du système, soit un modèle géocentrique.

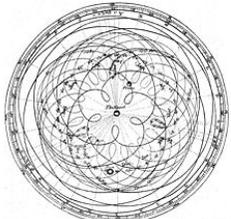


60

Claude Ptolémée, né vers 100 - mort vers 168 à Canope (Egypte)

### Modèle géocentrique

- Pour expliquer les connaissances acquises au cours des siècles suivants, le modèle nécessitait des séries importantes d'orbites autour des orbites d'ordre inférieur.




61

### Les sphères armillaires

- Ce modèle est utilisé pendant très longtemps et encore au 16<sup>ème</sup> siècle.

Jan Gossaert (ca 1478 – 1532), Portrait de jeune princesse portant une sphère armillaire, vers 1530, The National Gallery, Londres




62



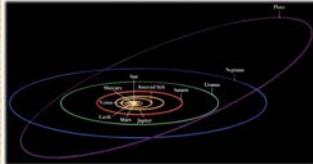
63

Galilée, né à Pise en 1564 et mort à Arcetri près de Florence le 8 janvier 1642 (77 ans)

### Modèle héliocentrique

- On va donc chercher utiliser un critère de meilleur ajustement mais en pénalisant par le nombre de paramètres utilisés dans le modèle pour arriver à la proposition de Copernic, reprise par Galilée, du modèle héliocentrique.

Nicolas Copernic est un astronome polonais, également chanoine, médecin et mathématicien, né le 19 février 1473 à Thorn (Toruń), Prusse royale (royaume de Pologne) et mort le 24 mai 1543 à Frauenburg Prusse royale (royaume de Pologne).





64

## Comparer des modèles

Soit  $m_1$  et  $m_2$ , deux modèles avec  $p_1$  et  $p_2$  paramètres ajustés sur un même jeu de données  $d$  et ayant pour vraisemblance  $L_1$  et  $L_2$ , on va définir la statistique AIC comme étant:

$$AIC = -2 \ln L + 2p$$

Akaike, H. 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control 19:716-723.

Entre deux modèles, on choisira donc l'AIC (Akaike Information Criterion) est le plus faible: meilleur ajustement mais pas trop de paramètres.

65

## Comparer des modèles

L'AIC n'est pas un test au sens de la théorie des tests.

On peut convertir l'AIC en une statistique qui représente la probabilité que ce soit réellement le meilleur modèle parmi ceux testés.

C'est l'Akaike weight:

$$\frac{\exp\left(-\frac{1}{2}(AIC_i - AIC_{min})\right)}{\sum_{j=1}^k \exp\left(-\frac{1}{2}(AIC_j - AIC_{min})\right)}$$

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York

66

## AIC, AICc, BIC

- Il existe d'autres critères que l'AIC.
  - L'AICc est une version corrigée de l'AIC pour les petits effectifs. Il évite une surparamétrisation.
  - Le BIC (Bayesian Information Criterion) peut-être utilisé lorsque le « vrai » modèle fait partie des modèles testés.

*There is no true model.*

Anderson & Burnham 1999

67

## BIC

- Dans des cas particuliers, on peut se trouver en présence du « vrai » modèle. Par exemple si le test est du genre:
  - 'C'est différent' ou 'c'est identique'.
    - Le « vrai » modèle est présent.

Girondot, M. & Guillon, J.-M. (2018) The w-value: An alternative to t- and  $X^2$  tests. *Journal of Biostatistics & Biometrics*, 1, 1-4.

68

### Distribution statistique des effets

- Les effets peuvent être des variables quantitatives continues ou bien des données qualitatives de type morte ou vivante.
- La distribution des données quantitatives ont comme caractéristiques d'être :
  - Continue
  - Positive
  - Généralement unimodale
  - Asymétrique
- La distribution des données qualitatives sont du type :
  - 0 - 1

69

### Distribution des données

- La distribution des données quantitatives ont comme caractéristiques d'être :
  - Continue
  - Positive
  - Généralement unimodale
  - Asymétrique
- La distribution des données qualitatives sont du type :
  - 0 - 1
- Revenons aux caractéristiques de la loi normale:
  - Non bornée
  - Symétrique

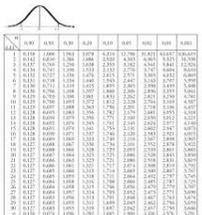
70

### Donc pourquoi continuer à utiliser une distribution que l'on sait non-appropriée ?

On a d'autres distributions qui sont flexibles et plus appropriées:

- La binomiale pour les données de type 0-1
- La binomiale négative pour les comptages
- La gamma pour les valeurs  $\geq 0$
- La lognormale pour les valeurs  $> 0$

Cela était compréhensible à une époque, pas si lointaine, où les statistiques utilisaient des tables papiers...



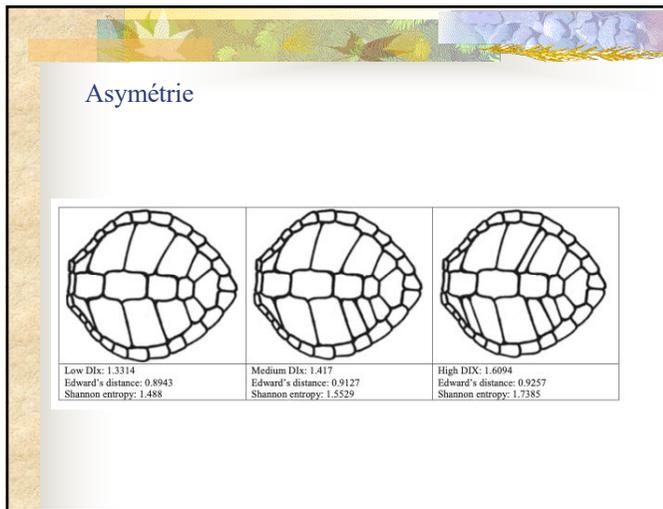
71

### Exemple

Cortés-Gómes, A. A., et al. (2018). "Carapace asymmetry: A possible biomarker for accumulation in adult Olive Ridley marine turtles?" Marine Pollution Bulletin 129: 92-101.




72

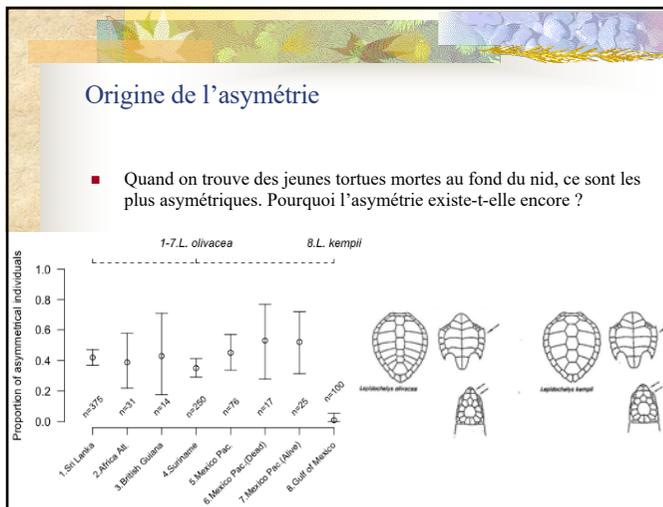


73

### Les données

- 17 individus femelle sur lesquels ont été mesurés:
  - L'asymétrie de la carapace
  - La concentration en Al, As, B, Cd, Cr, Cu, Li, Mn, Pb, Se, Sr, Ti, Zn dans "Albumin" "Blood" "Bone" "Brain" "Kidney" "Liver" "Muscle" "Yolk"

74



75

### Métaux et métalloïdes ou ETM

- La définition des métaux lourds n'est pas basée sur la chimie mais sur un concept industriel. Ils sont parfois définis comme les éléments métalliques ayant une masse volumique supérieure à 5000 kg/m<sup>3</sup>, mais ce seuil est parfois ramené à 4000 kg/m<sup>3</sup>.
- L'Europe a proposé une définition retenue pour le droit européen et celui des Etats-membres : «un métal lourd désigne tout composé d'antimoine, d'arsenic, de cadmium, de chrome hexavalent, de cuivre, de plomb, de mercure, de nickel, de sélénium, de tellure, de thallium et d'étain, ainsi que ces matériaux sous forme métallique, pour autant qu'ils soient classés comme substances dangereuses».

76

## Relation entre asymétrie et concentration en éléments traces

- On établit un modèle linéaire liant l'indice d'asymétrie et la concentration en éléments traces:
- $a \text{ Al} + b \text{ As} + c \text{ B} + d \text{ Cd} + e \text{ Cr} + f \text{ Cu} + g \text{ Li} + h \text{ Mn} + i \text{ Pb} + j \text{ Se} + k \text{ Sr} + l \text{ Ti} + m \text{ Zn} + \text{Cst}$
- On cherche les valeurs de  $a \dots m$  et  $\text{Cst}$  qui maximisent la vraisemblance des observations d'asymétrie dans ce modèle avec l'hypothèse que la distribution de l'asymétrie est normale.
- On peut établir l'erreur standard de chacun des estimateurs  $a \dots m$  et  $\text{Cst}$ , c'est-à-dire une mesure de la qualité de l'estimation de la valeur de chacun des paramètres.
  - *D'un point de vue technique, on utilise la matrice hessienne comme une approximation de l'inverse de la matrice des variances-covariances.*

77

## Résultat

Coefficients:

|             | Estimate   | Std. Error | t value | Pr(> t )   |
|-------------|------------|------------|---------|------------|
| (Intercept) | 1.551e+00  | 2.198e-02  | 70.544  | <2e-16 *** |
| Al          | -4.686e-05 | 3.241e-04  | -0.145  | 0.8853     |
| As          | -5.181e-03 | 3.022e-03  | -1.714  | 0.0895 .   |
| B           | -4.013e-03 | 2.467e-03  | -1.627  | 0.1068     |
| Cd          | 5.540e-04  | 2.362e-04  | 2.346   | 0.0209 *   |
| Cr          | 3.028e-02  | 1.561e-02  | 1.940   | 0.0551 .   |
| Cu          | -1.085e-03 | 2.490e-03  | -0.436  | 0.6639     |
| Li          | 5.339e-02  | 3.549e-02  | 1.504   | 0.1356     |
| Mn          | 6.966e-05  | 3.698e-04  | 0.188   | 0.8509     |
| Pb          | -9.036e-03 | 7.237e-02  | -0.125  | 0.9009     |
| Se          | 5.932e-03  | 3.984e-03  | 1.489   | 0.1396     |
| Sr          | 6.279e-06  | 1.474e-04  | 0.043   | 0.9661     |
| Ti          | -9.248e-04 | 2.880e-03  | -0.321  | 0.7487     |
| Zn          | -7.157e-04 | 6.513e-04  | -1.099  | 0.2743     |

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

78

## Qu'est ce que le $\text{Pr}(>|t|)$ ?

- Il s'agit d'une statistique qui est la p-value, centrale depuis les années 1920-1930 dans les tests statistiques.

79

## Définitions verbales

- Dans un test de type  $H_0-H_1$ , l'hypothèse  $H_0$  est dite hypothèse nulle. Elle est clairement explicitée. L'hypothèse  $H_1$  est l'hypothèse alternative. Ce sont toutes les hypothèses qui ne sont pas  $H_0$ .
- Quand on fait un test, on cherche si le jeu de données a pu être obtenu sous l'hypothèse  $H_0$ . Il faut noter que la réponse est tout le temps « oui » mais avec une probabilité plus ou moins élevée. Le « p value » désigne cette probabilité (Sir Ronald Fisher, Statistical Methods for Research Workers, 1925).

$H_0$ : Hypothèse nulle      x: jeu de données  
 $H_1$ : Hypothèse alternative      p: probabilité que le jeu de données x ait pu être obtenu avec  $H_0$

80

## Définitions verbales

- Donc le « p value » représente la probabilité de se tromper si on rejette  $H_0$ . Par exemple, si  $p=0,2$ , cela signifie que si on rejette  $H_0$ , on sait que ce jeu de données avait 20% de chance d'être obtenu alors que  $H_0$  était vraie. Donc on prend un risque assez grand en rejetant  $H_0$ . Il faut dire qu'on ne rejette pas  $H_0$ .
- *Attention: on n'accepte jamais  $H_0$  car on ne peut jamais rejeter  $H_1$*

81

## Les valeurs de p

- Si un effet possède un niveau  $p$  de 0,019, et qu'un autre possède un niveau  $p$  de 0,048, il peut être incorrect de conclure que nous avons la preuve statistique que le premier effet est plus fort que le second.
- Pour interpréter correctement ces résultats, nous avons besoin de plus d'information. Pour comprendre pourquoi, supposons qu'un article reporte une valeur  $p$  de 0,001. Ce niveau  $p$  peut être lié à un effet évident dans la population avec une taille d'échantillon importante, ou un effet important dans la population avec une taille d'échantillon modérée, ou encore avec un effet très important dans la population avec une petite taille d'échantillon.
- De la même manière, un niveau  $p$  de 0,075 peut représenter un effet important combiné à une taille d'échantillon faible, ou un effet minuscule avec un échantillon de très grande taille.
- Il faut donc prendre toutes les précautions nécessaires pour comparer des niveaux  $p$ .

82

## Définition

- La p-value est donc, d'un point de vue formel:  $\text{prob}(x|H_0)$ , le signe | signifiant « sachant »;
- C'est donc la probabilité d'observer les données  $x$  sachant que  $H_0$  est vrai.
- *Mais est-ce vraiment ce qu'on veut savoir ?*

83

## On cherche en fait...

- $\text{prob}(x|H_0)$  ne nous donne pas la probabilité que  $H_0$  soit vraie mais la probabilité que les données  $x$  aient pu avoir été obtenues sous  $H_0$ .
- Or on est intéressé par  $\text{prob}(H_0|x)$  ce qui se lit: quelle est la probabilité de  $H_0$  sachant qu'on a observé les données  $x$ .

Hubbard, R., Lindsay, R.M., 2008. Why P values are not a useful measure of evidence in statistical significance testing. *Theory & Psychology* 18, 69-88.

84

Thomas Bayes (1702-1761) 

**Théorème de Bayes**

- $P(A|B) P(B) = P(A, B)$

Par symétrie:

- $P(B|A) P(A) = P(A, B)$

85

**D'après le théorème de Bayes**

- $\text{prob}(H_0, x) = \text{prob}(H_0) \text{prob}(x|H_0)$   
et
- $\text{prob}(H_0, x) = \text{prob}(x) \text{prob}(H_0|x)$
- Donc:  $\text{prob}(H_0|x) = \text{prob}(H_0, x)/\text{prob}(x)$   
 $\text{prob}(H_0|x) = \text{prob}(x|H_0) \text{prob}(H_0) / \text{prob}(x)$
- **Comme**  $\text{prob}(H_0) \neq \text{prob}(x)$
- **Donc:  $\text{prob}(H_0|x) \neq \text{prob}(x|H_0)$**

86

**prob( $H_0|x$ ) vs. prob( $x|H_0$ )**

- En utilisant un test statistique Bayésien sur une moyenne normale, James Berger et Thomas Sellke ont montré que pour des valeurs de  $p$  de 0,05, 0,01, et 0,001, les probabilités postérieures du modèle nul,  $\text{prob}(H_0|x)$ , pour  $n = 50$  sont 0,52, 0,22, et 0,034. Pour  $n = 100$  les valeurs étaient de 0,60, 0,27, et 0,045.

Berger, J.O., Sellke, T., 1987. Testing a point null hypothesis: The irreconcilability of  $p$  values and evidence (with comments). *Journal of the American Statistical Association* 82, 112-139.

87

**$p = 0.05$**

- "If you use  $p = 0.05$  to suggest that you have made a discovery, you will be wrong at least 30% of the time."

Colquhoun, D., 2014. An investigation of the false discovery rate and the misinterpretation of  $p$ -values. *Royal Society Open Science* 1, 140216-140216.

88

## Nature...

- Nuzzo RL (2014) Scientific method: Statistical errors. Nature 506:150-152
  - P values, the 'gold standard' of statistical validity, are not as reliable as many scientists assume.

89

## American Statistician Association

- Baker M (2016) Statisticians issue warning on P values. Nature 351:151-152
- Wasserstein RL, Lazar NA (2016) The ASA's statement on p-values: context, process, and purpose. The American Statistician 70:129-133
  - Q: Why do so many colleges and grad schools teach  $p = .05$ ?
  - A: Because that's still what the scientific community and journal editors use.
  - Q: Why do so many people still use  $p = 0.05$ ?
  - A: Because that's what they were taught in college or grad school.

90

## Erreurs classiques et mauvaises interprétations sur les p-values

- La p-value n'est pas la probabilité que l'hypothèse nulle soit vraie, ni la probabilité que l'hypothèse alternative soit fausse
- La p-value n'est pas la probabilité que les données soient dues simplement à un coup de chance.
- La p-value n'est pas la probabilité de rejeter à tort l'hypothèse nulle.
- La p-value n'est pas la probabilité que reproduire l'expérience donnerait la même conclusion.
- Le seuil de signification, par exemple 0,05, n'est pas déterminé par la valeur de p.
- La p-value ne représente pas la force ou l'importance de l'effet observé.

91

## Pourquoi maintenant ?

- The Structure of Scientific Revolutions est un essai rédigé par le philosophe et historien des sciences, Thomas Samuel Kuhn (1962).
- Il y démontre notamment comment un paradigme en vient à en remplacer un autre.

92

## Alors, que faire ?

- Comparaison des intervalles de confiance issus de modèles bayésiens
- Construction de modèles et comparaison des résultats par des méthodes issues des théories de l'information (AIC, AICc)

93

## On revient à notre résultat...

```
Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.551e+00 2.198e-02 70.544 <2e-16 ***
Al -4.686e-05 3.241e-04 -0.145 0.8853
As -5.181e-03 3.022e-03 -1.714 0.0895 .
B -4.013e-03 2.467e-03 -1.627 0.1068
Cd 5.540e-04 2.362e-04 2.346 0.0209 *
Cr 3.028e-02 1.561e-02 1.940 0.0551 .
Cu -1.085e-03 2.490e-03 -0.436 0.6639
Li 5.339e-02 3.549e-02 1.504 0.1356
Mn 6.966e-05 3.698e-04 0.188 0.8509
Pb -9.036e-03 7.237e-02 -0.125 0.9009
Se 5.932e-03 3.984e-03 1.489 0.1396
Sr 6.279e-06 1.474e-04 0.043 0.9661
Ti -9.248e-04 2.880e-03 -0.321 0.7487
Zn -7.157e-04 6.513e-04 -1.099 0.2743
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Donc plutôt que d'utiliser la p-value, on va utiliser l'AIC.

94

## Relation entre asymétrie et concentration en éléments traces

- On établit un modèle linéaire liant l'indice d'asymétrie et la concentration en éléments traces:
  - a Al + ... + d Cd + ... + Cst qui donne une vraisemblance de  $L_1$
  - a Al + ... + ... + Cst qui donne une vraisemblance de  $L_2$
- La différence entre les deux modèles est l'inclusion ou non de B
- On peut donc calculer l'AIC des deux modèles par
  - $AIC_1 = -2 \ln L_1 + 2 \times 13$  et  $AIC_2 = -2 \ln L_2 + 2 \times 12$
- On peut alors savoir si l'inclusion de B améliore ou non la capacité à expliquer l'asymétrie.

95

## Résultat

```
> g1 <- glm(formula = d11 ~ Al + As + B + Cd + Cr
+ Cu + Li + Mn + Pb + Se + Sr + Ti + Zn , data =
newassymetry)
> g2 <- glm(formula = d11 ~ Al + As + B + Cr + Cu
+ Li + Mn + Pb + Se + Sr + Ti + Zn , data =
newassymetry)
> compare_AIC(Model1=g1, Model2=g2)
      AIC DeltaAIC Akaike_weight
Model1 -155.4951 0.000000      0.8853828
Model2 -151.4062 4.088844      0.1146172
```

## Rappel

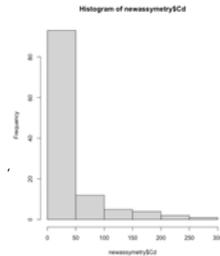
```
Cd      5.540e-04 2.362e-04 2.346 0.0209 *
```

96

## Mais la distribution normale est-elle justifiée ?

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.551e+00  2.198e-02  70.544  <2e-16 ***
Al           -4.686e-05  3.241e-04  -0.145  0.8853
As           -5.181e-03  3.022e-03  -1.714  0.0895 .
B            -4.013e-03  2.467e-03  -1.627  0.1068
Cd           5.540e-04  2.362e-04  2.346  0.0209 *
Cr           3.028e-02  1.561e-02  1.940  0.0551 .
Cu          -1.085e-03  2.490e-03  -0.436  0.6639
Li           5.339e-02  3.549e-02  1.504  0.1356
Mn           6.966e-05  3.698e-04  0.188  0.8509
Pb          -9.036e-03  7.237e-02  -0.125  0.9009
Se           5.932e-03  3.984e-03  1.489  0.1396
Sr           6.279e-06  1.474e-04  0.043  0.9661
Ti          -9.248e-04  2.880e-03  -0.321  0.7487
Zn          -7.157e-04  6.513e-04  -1.099  0.2743
    
```



```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

97

## Utilisation d'une distribution Gamma

```

              Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.444e-01  9.168e-03  70.291  <2e-16 ***
Al           2.204e-05  1.291e-04  0.171  0.8648
As           2.096e-03  1.261e-03  1.662  0.0995 .
B            1.668e-03  1.000e-03  1.668  0.0993 .
Cd          -2.254e-04  9.480e-05  -2.378  0.0193 *
Cr          -1.261e-02  6.349e-03  -1.986  0.0497 *
Cu           4.405e-04  1.034e-03  0.426  0.6710
Li          -2.125e-02  1.407e-02  -1.510  0.1341
Mn          -2.703e-05  1.535e-04  -0.176  0.8606
Pb           6.051e-03  3.007e-02  0.201  0.8409
Se          -2.430e-03  1.619e-03  -1.500  0.1366
Sr          -7.375e-06  5.967e-05  -0.124  0.9019
Ti           3.510e-04  1.148e-03  0.306  0.7603
Zn           3.144e-04  2.703e-04  1.163  0.2475
    
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

              AIC DeltaAIC Akaike_weight
Modell1 -150.2030 0.000000  0.8768582
Modell2 -146.2769 3.926017  0.1231418
    
```

98

## Mais...

- Ce n'est pas le seul problème !
- On a des mesures qui proviennent d'organes différents. Il faut le prendre en compte car cela peut générer des corrélations entre données.
- On utilisera alors un GLMM (le deuxième M = mixte avec des effets aléatoires et des effets fixes).

99

## Résultat

```

              Estimate Std. Error t value Pr(>|z|)
(Intercept)  6.444e-01  8.408e-03  76.645  < 2e-16 ***
Al           2.204e-05  1.033e-04  0.213  0.83099
As           2.096e-03  1.187e-03  1.766  0.07743 .
B            1.668e-03  9.313e-04  1.791  0.07323 .
Cd          -2.254e-04  8.609e-05  -2.619  0.00883 **
Cr          -1.261e-02  6.016e-03  -2.096  0.03612 *
Cu           4.405e-04  9.588e-04  0.459  0.64594
Li          -2.125e-02  7.856e-03  -2.705  0.00684 **
Mn          -2.703e-05  1.456e-04  -0.186  0.85272
Pb           6.051e-03  2.578e-02  0.235  0.81445
Se          -2.430e-03  1.479e-03  -1.643  0.10040
Sr          -7.375e-06  1.637e-05  -0.450  0.65240
Ti           3.510e-04  9.315e-04  0.377  0.70631
Zn           3.144e-04  1.837e-04  1.711  0.08706 .
    
```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

100



Mais...

- Il n'existe pas d'AIC pour les modèles linéaires généralisés mixtes !

101