

P.-O. Caron
Université TÉLUQ
SQRP 2025

Méthodes quantitatives:
Sortir des sentiers battus!

S'hisser soi-même par le tirant de ses bottes



Bootstrap



Pull yourself up by your bootstraps - Rudolf Erich Raspe

S'hisser soi-même par le tirant de ses bottes
(traduction libre)

- Se débrouiller seul
- Fût un temps où les nouvelles méthodes avec des noms originaux (Monte Carlo, forêt aléatoire, Métropolis, Kolmogorov-Smirnov)

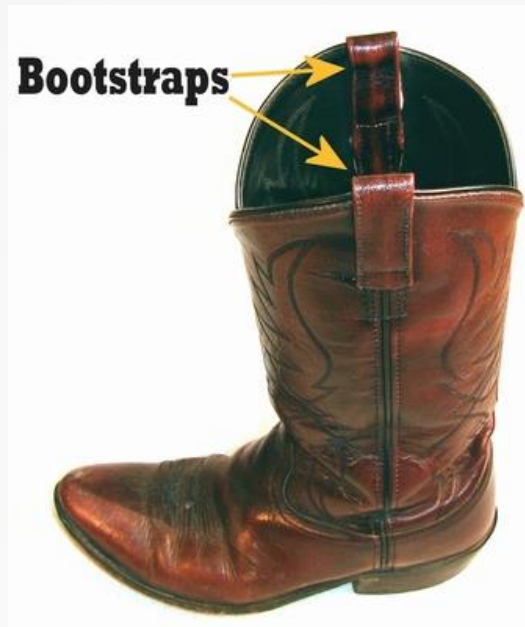
BOOTSTRAPS

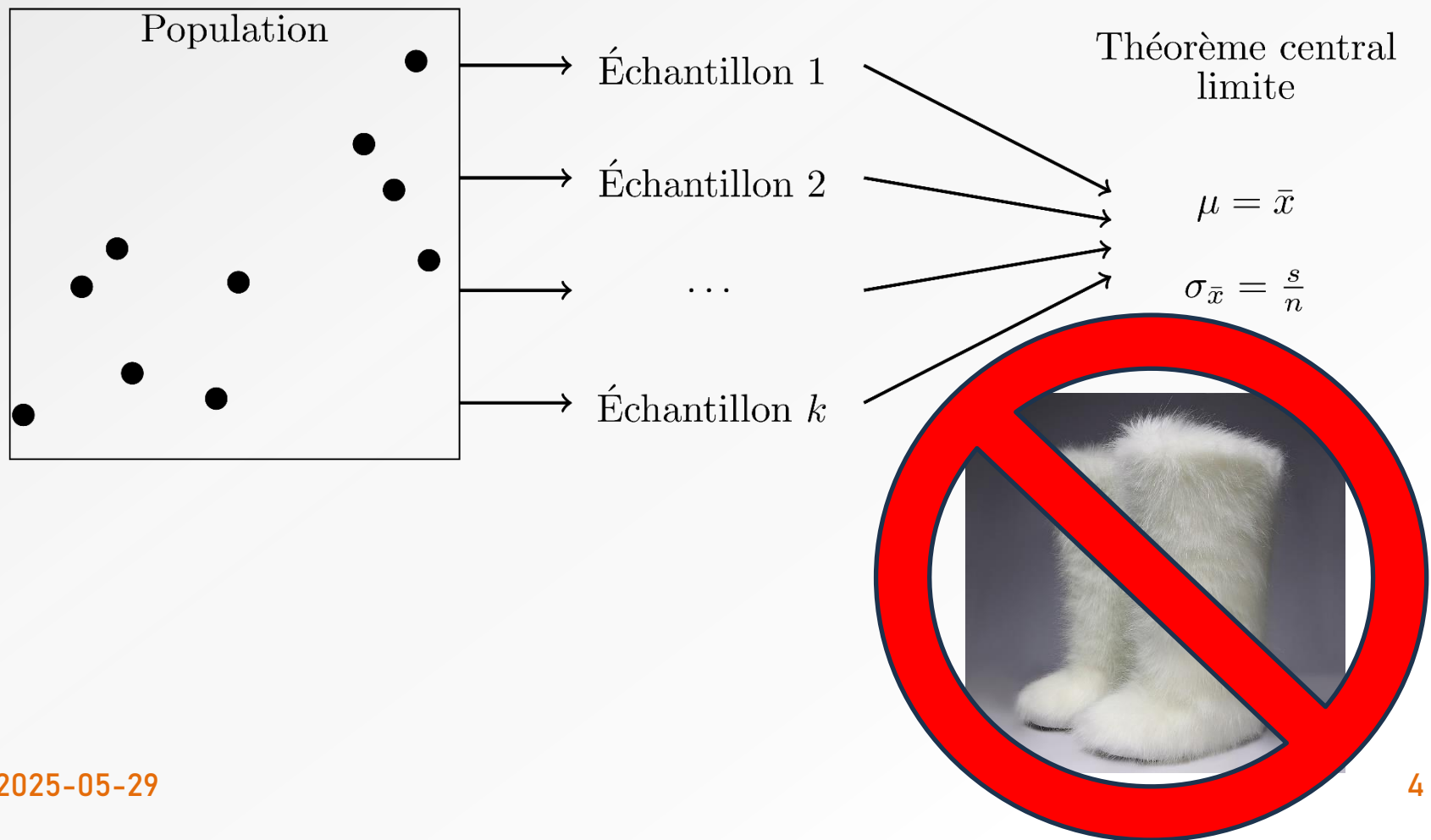


Bootstrap



- Famille des simulations Monte-Carlo
- Méthode statistique incontournable
- Implémentation à l'ordinateur de concepts statistiques





Bootstrap



- Quand la distribution d'échantillonnage n'est pas normale
- Quand l'erreur standard est difficile à estimer
- *Quand la taille d'échantillon est petite*



Bootstrap



- Permet
 - d'estimer l'indice désiré
 - d'estimer son erreur standard
 - d'estimer ses intervalles de confiance
 - de réaliser un test d'hypothèse
- Tout cela en ignorant (presque tous) les postulats sous-jacents traditionnels



Bootstrap

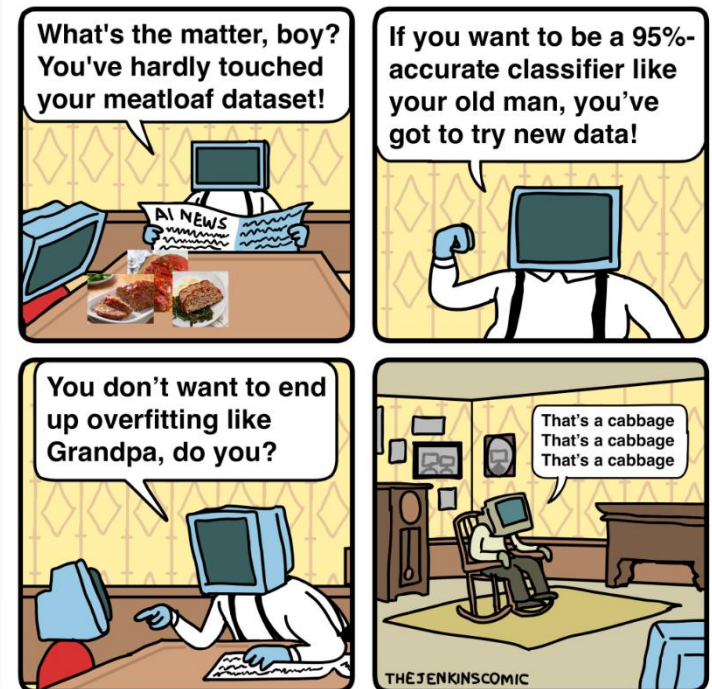


- Deux hypothèses fondamentales :
 - l'échantillon reflète la population
 - chaque unité est indépendante et identiquement distribuée

Bootstrap



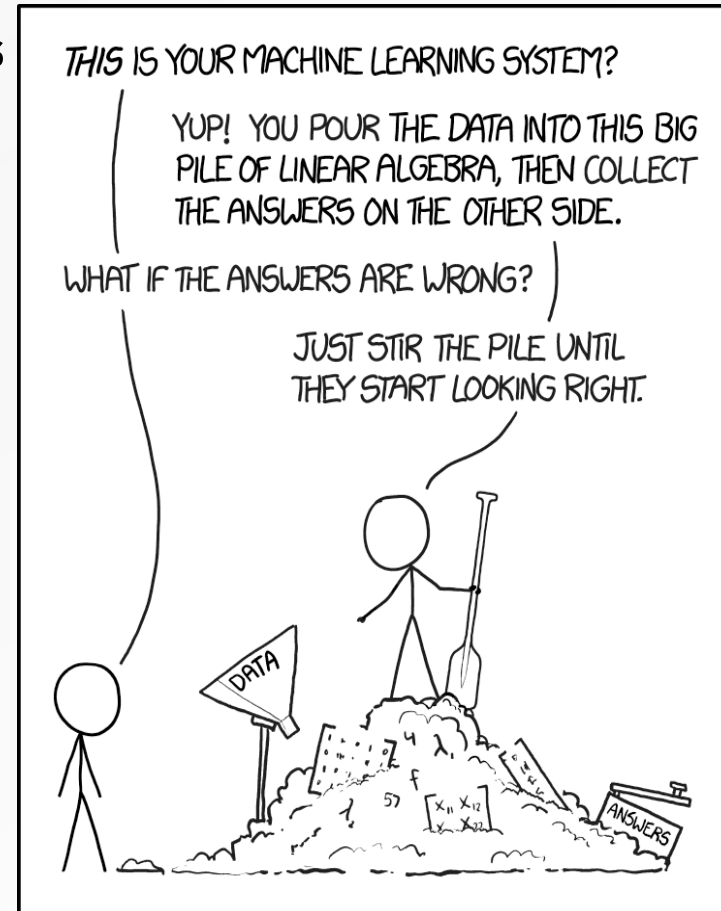
- L'échantillon initial est considéré comme une pseudopopulation
- Aucune information autre que celle fournie par l'échantillon est nécessaire
- Donne une distribution d'échantillonnage

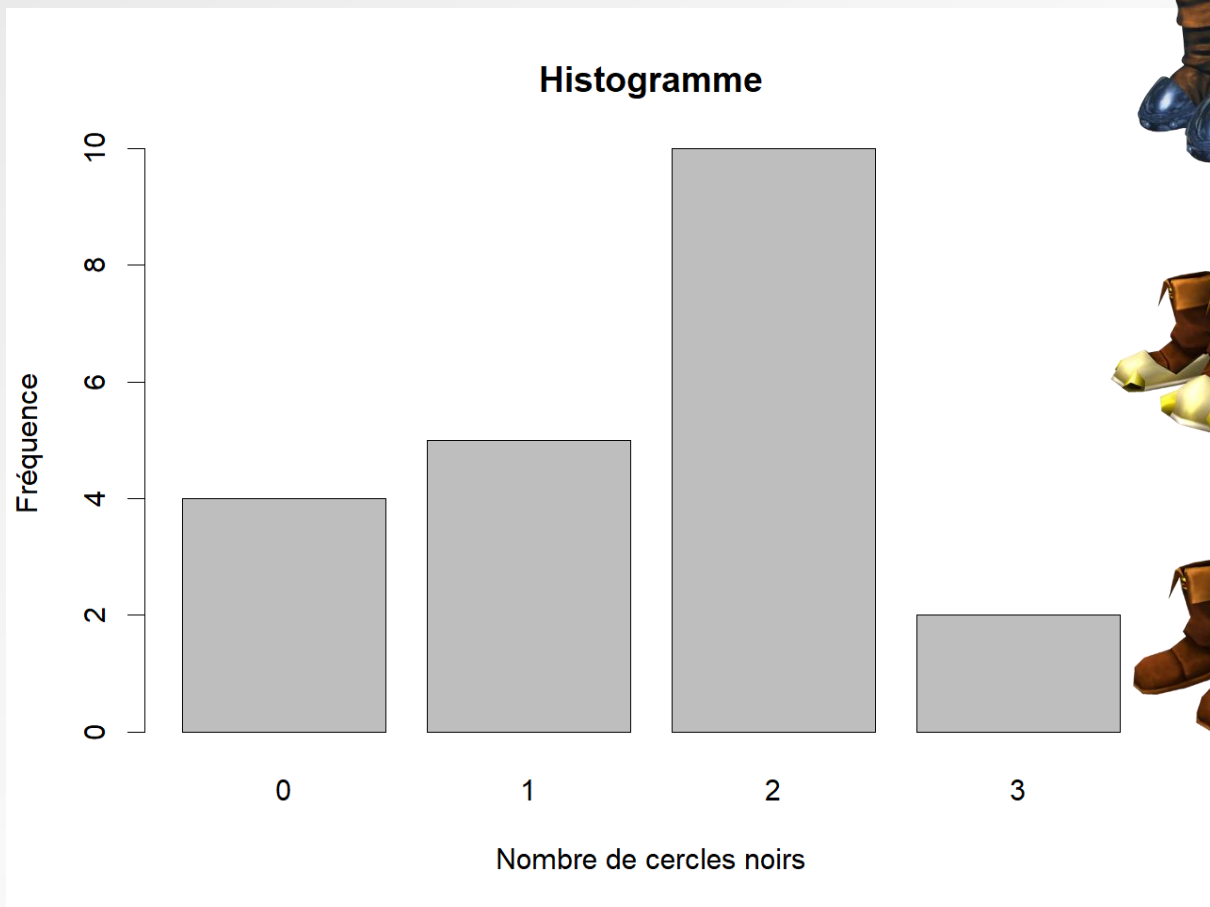


Étapes



- Construit à partir d'un vrai jeu de données
- Repose sur les étapes suivantes :
 1. Sélectionner avec remise les unités d'un échantillon;
 2. Calculer et enregistrer l'indice statistique désiré auprès de ce nouvel échantillon;
 3. Réitérer les deux premières étapes un nombre élevé de fois.





Échantillon
 $\bar{x} = .666$
 $s_x = .272$



Bootstrap (1000)
 $\bar{x} = .660$
 $s_x = .280$



Bootstrap (21)
 $\bar{x} = .523$
 $s_x = .326$

R



```
# Le jeu de données D
ns <- nrow(D)           # Nombre d'unités
nreps <- 1000           # Le nombre de rééchantillonnage
indice <- numeric()     # Variable vide pour l'enregistrement

# L'utilisation d'une boucle pour réitérer les étapes
for (i in 1:nreps){
  # Rééchantillonnage avec remise
  idx <- sample(ns, replace = TRUE)

  # Nouvel échantillon
  nouveau.D <- D[idx,]

  # Calculer et enregistrer l'indice
  indice[i] <- calcul_indice(nouveau.D)
}
```



```
int getRandomNumber()
{
  return 4; // chosen by fair dice roll.
            // guaranteed to be random.
}
```

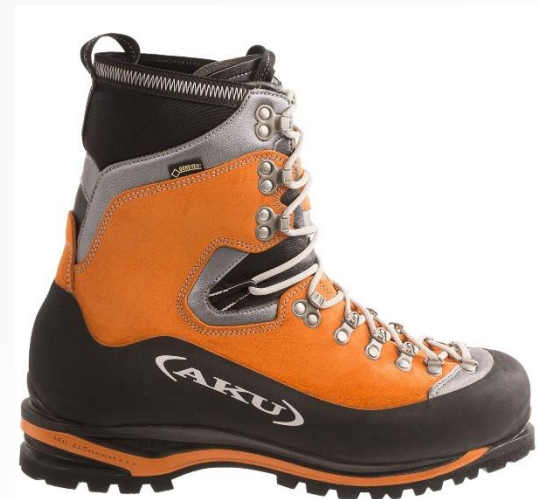
R



```
# Histogramme  
hist(indice)
```

```
# Tendances centrales  
mean(indice); median(indice)
```

```
# Intervalles de confiance  
alpha <- .05  
quantile(indice, probs = c(alpha/2, (1-alpha)/2))
```



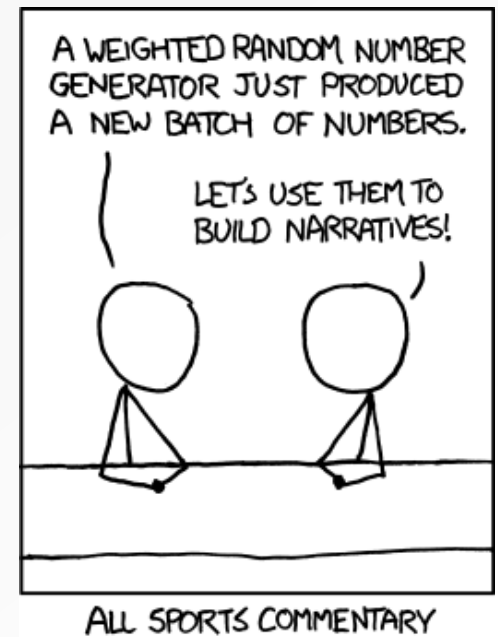
R



```
echan <- c(0,1,1)
nreps <- 1000
ind <- numeric()

for(i in 1:nreps){
  idx <- sample(length(echan), replace = TRUE)
  ind[i] <- mean(echan[idx])
}
```

```
mean(ind)
sd(ind)
```



Les indices



- Médiane
- Coefficient de détermination
- Coefficient de régression
- Taille d'effet
- Analyse de médiation
- Analyse multiniveau
- Modèle d'équations structurelles



Faiblesses



- Pas une panacée
 - Le bootstrap est meilleur pour estimer des tendances centrales que des extrémités
 - Peut être biaisé
 - P.ex. les totaux, les intervalles de confiance
 - Dépend de l'estimateur
 - Le bootstrap naïf peut être inconsistant ou asymptotiquement invalide
- R n'est pas particulièrement efficace pour réaliser des boucles



Versions



- Basic bootstrap
- Parametric bootstrap
- Non parametric bootstrap
- Bias-corrected and accelerated (BCa)
- Bayesian bootstrap



Package R



- bootstrap (Tibshirani & Leisch, 2019)
- boot (Canty & Ripley, 2021).
- Plusieurs fonctions R intègrent le bootstrap
- Facile à implante soi-même





FÉLICITATIONS!
(et merci de votre attention)

Références



- Canty, A., & Ripley, B. D. (2021). *boot: Bootstrap R (S-Plus) functions*.
- Caron, P.-O. (2023). *Méthodes quantitatives avec R*.
<https://mqr.telug.ca>
- Efron, B., & Tibshirani, R. (1979). *An introduction to the bootstrap*. Chapman & Hall.
- Tibshirani, R., & Leisch, F. (2019). *bootstrap: Functions or the book "An Introduction to the bootstrap"*.
<https://CRAN.R-project.org/package=bootstrap>

