

# Séance 4 : Loi de Pareto et extrapolations

**Professeur :** Christophe ANCEY

**Assistants :** Clemente GOTELLI, Mehrdad KIANI

## Ajustements et extrapolations

Nous allons comparer ici la pluie centennale à Davos obtenue par extrapolation grâce aux quatre méthodes vues en cours, à savoir :

- **Méthodes à seuils :** Méthode du renouvellement et loi généralisée de Pareto
- **Méthodes par blocs :** Méthodes des moments et du maximum de vraisemblance

## 1 Méthode du renouvellement

On rappelle que la méthode du renouvellement cherche à représenter les événements extrêmes grâce à une loi de Poisson (comptabilisant le nombre d'événements pour une période donnée) et une loi exponentielle (traduisant la distribution de ces événements dans le temps).

Pour une période de retour  $T$  très grande, on obtient le quantile en fonction de la période de retour grâce à la formule suivante :

$$C = s + \frac{\ln \lambda}{\mu} + \frac{1}{\mu} \ln T + O(T^{-1}) \quad (1)$$

Avec les paramètres :

- $\lambda = n_s / n_a$
- $\mu = 1 / (\bar{C} - s)$

1. Tracer l'histogramme des précipitations à Davos couvrant  $n_a$  années et déterminer à priori une valeur  $s$  du seuil possible. Ne retenir que les  $n_s$  précipitations dépassant le seuil par la suite.
2. Classer les valeurs de l'échantillon par ordre croissant et calculer  $\bar{C}$ , sa moyenne empirique.
3. A chaque valeur de l'échantillon de rang  $i$ , on affecte la période de retour :

$$T_i = \frac{n_a}{n_s} \frac{n_s + 1}{i} \quad (2)$$

4. Calculer les paramètres  $\lambda$  et  $\mu$ , puis tracer dans un diagramme  $(T, C)$  la variation du quantile  $C$  en fonction de la période de retour  $T$ , grâce à la l'équation 1. Comparer avec les valeurs empiriques.
5. Quel est la valeur de la pluie centennale ?  $C_{100} =$
6. Changer le seuil tel que  $s_2 > s$  puis  $s_3 < s$ , et tracer les lois obtenues sur la même fenêtre. Quelle influence cela a-t-il sur l'extrapolation ? Comment choisir un « bon » seuil selon vous ?

## 2 Loi de Pareto

*Vilfredo Pareto, né le 15 juillet 1848 à Paris et mort le 19 août 1923 à Céligny (Suisse), est un sociologue et économiste italien. Il demeure célèbre pour son observation des 20 % de la population qui possèdent 80 % des richesses en Italie, généralisée plus tard en distribution de Pareto. Cette observation a été étendue à d'autres domaines sous le terme de « principe de Pareto ». Il définit la notion d'optimum paretien comme une situation d'ensemble dans laquelle un individu ne peut améliorer sa situation sans détériorer celle d'un autre individu.*

La loi de Pareto s'exprime ainsi :

$$G(x) = 1 - \left( 1 + \frac{\tilde{\xi} x}{\tilde{\sigma}} \right)^{-1/\tilde{\xi}} \quad (3)$$

On cherche à déterminer ses paramètres. On peut obtenir que pour tout seuil  $s > s_0$ , l'équation suivante est vérifiée :

$$E[X - s | X > s] = \frac{\tilde{\sigma}_s}{1 - \tilde{\xi}} = \frac{\tilde{\sigma}_{s_0} + \tilde{\xi}(s - s_0)}{1 - \tilde{\xi}} \quad (4)$$

$\tilde{\xi}/(1 - \tilde{\xi})$  est donc la pente et  $(\tilde{\sigma}_{s_0} - \tilde{\xi}s_0)/(1 - \tilde{\xi})$  l'ordonnée à l'origine de la fonction  $f(s) = E[X - s | X > s]$

1. Tracer  $E(X - s | X > s)$  en fonction de  $s$  et déterminer a priori sur quel domaine le paramètre  $\tilde{\xi}$  pourra être ajusté. (calculer la moyenne de la distance des valeurs qui dépassent le seuil par rapport à ce même seuil)
2. Pour un domaine fixé, calculer les paramètres de la loi de Pareto qui approchent le mieux cette distribution (vous pouvez utiliser la commande `polyfit`). Tracer la droite ajustée sur le même graphique que  $f(s)$ . A quelle type de loi de valeurs extrêmes peut-on comparer la loi de Pareto ?
3. Tracer le quantile  $x_p(T)$  ainsi ajusté et comparer le aux données de l'échantillon.

$$\begin{aligned} \text{si } \tilde{\xi} \neq 0, x_p(T) &= s + \frac{\tilde{\sigma}}{\tilde{\xi}} \left( \left( T \frac{n_s}{n_a} \right)^{\tilde{\xi}} - 1 \right) \\ \text{si } \tilde{\xi} = 0, x_p(T) &= s + \tilde{\sigma} \ln \left( T \frac{n_s}{n_a} \right) \end{aligned} \tag{5}$$

4. Ajuster maintenant les paramètres de la loi de Pareto grâce à la méthode d'optimisation de votre choix pour 3 seuils différents (max. de vraisemblance, Hasting-Metropolis, méthode des moments, etc) [`gpfitt`]
5. Reproduire ces résultats sur un même graphique et les comparer.
6. Donner la valeur du quantile extrapolé  $C_{100} =$

### 3 Maximum de vraisemblance

1. Calculer les maxima annuels.
2. Ajuster une loi de valeurs extrêmes grâce au maximum de vraisemblance.
3. Tracer sur la même figure, le quantile obtenu  $x_{ev}(T)$  via la méthode par bloc et les quantiles  $x_p(T)$  et  $C_s(T)$  obtenus via les méthodes à seuil. Conclure.