Prédire les extrêmes

Le théorème central limite pour les maxima

Stéphane Lhaut

UCLouvain - Institut de Statistique, Biostatistique et Sciences Actuarielles

Brussel Summer School of Mathematics – 01/09/2023

UCLouvain

LIDAM

ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

fnrs
LA LIBERTÉ DE CHERCHER

Résumé

Inondation à Port Pirie

- Port Pirie est une petite ville du sud de l'Australie située dans l'enclave du golf Spencer.

- La ville a connu une grave inondation en 1934 à cause d'une hausse du niveau de la mer.

PORT PIRIE FLOODED

Two Drowned; 1000 Homeless

SEA WALL BREAKS

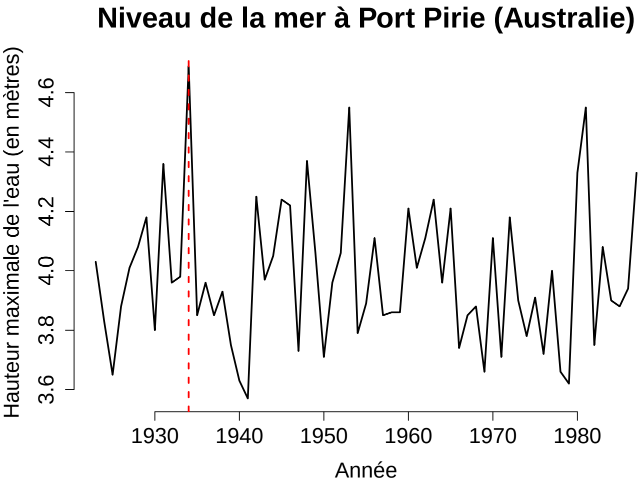
ADELAIDE, August 15.

Two lives have been lost and at least 1000 persons are homeless, and damage running into many thousands of pounds has been caused by tidal waters which swept into Port Pirie when an embankment gave way last night and inundated the greater part of the township.

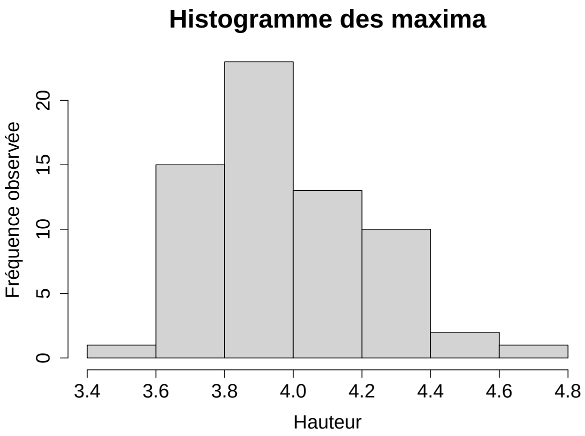
- Comment protéger la ville contre de potentielles futures montées des eaux ?

Maxima annuels du niveau de l'eau

- Le jeu de données `portpirie` du package R `evd` contient les maxima annuels du niveau de l'eau mesurés à Port Pirie en mètres au-dessus du niveau de la mer durant les années 1923–1987 (65 observations).



Dépasser les valeurs observées



- Comment quantifier la probabilité d'évènements encore plus extrêmes que ceux déjà observés par le passé ?

Convergence en distribution

Convergence faible

Une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ converge en distribution vers une variable aléatoire réelle X, ce qui sera noté $X_n \xrightarrow{D} X$, si

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) = F(x),$$

pour tout $x \in \mathbb{R}$ tel que F est continue en x.

Convergence en distribution

Convergence faible

Une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ converge en distribution vers une variable aléatoire réelle X, ce qui sera noté $X_n \xrightarrow{D} X$, si

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) = F(x),$$

pour tout $x \in \mathbb{R}$ tel que F est continue en x.

<u>Exemple :</u>

Si les $(Y_i)_{i\in\mathbb{N}}$ sont i.i.d. et ont comme fonction de répartition commune

$$H(x) = \frac{x}{x+1} I_{[0,\infty)}(x),$$

quelle est la distribution asymptotique de $X_n = \max(Y_1, \ldots, Y_n)/n$?

Convergence en distribution

Convergence faible

Une suite de variables aléatoires réelles $(X_n)_{n\in\mathbb{N}}$ converge en distribution vers une variable aléatoire réelle X, ce qui sera noté $X_n \xrightarrow{D} X$, si

$$\lim_{n\to\infty} F_n(x) = \lim_{n\to\infty} \mathbb{P}(X_n \leq x) = \mathbb{P}(X \leq x) = F(x),$$

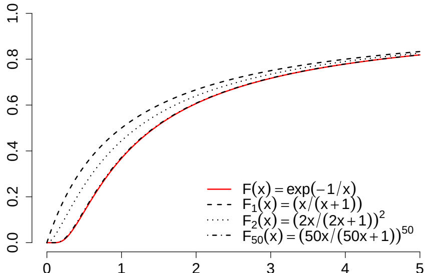
pour tout $x \in \mathbb{R}$ tel que F est continue en x.

<u>Exemple :</u>

Si les $(Y_i)_{i\in\mathbb{N}}$ sont i.i.d. et ont comme fonction de répartition commune

$$H(x) = \frac{x}{x+1} I_{[0,\infty)}(x),$$

quelle est la distribution asymptotique de $X_n = \max(Y_1, \ldots, Y_n)/n$?



Sommes de variables aléatoires i.i.d.

- Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles i.i.d.
- Notons $S_n = \sum_{i=1}^n X_i$.
- Le suite $(S_n)_{n\in\mathbb{N}}$ ne peut converger en distribution vers une distribution non-dégénérée puisque, par la loi des grands nombres (pourvu que $\mathbb{E}[|X_1|] < \infty$),

$$\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{D} \mathbb{E}[X_1].$$

- Si l'on souhaite obtenir une limite non dégénérée pour la suite $(S_n)_{n\in\mathbb{N}}$, il est donc nécessaire de la normaliser.

Théorème central limite

Théorème central limite (cas réel)

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles i.i.d. telle que

$$\mathbb{E}[|X_1|] < \infty \qquad \text{et} \qquad \mathbb{E}[X_1^2] < \infty.$$

La suite

$$\left(Z_n = \frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n(\mathbb{E}[X_1^2] - \mathbb{E}[X_1]^2)}}\right)_{n\in\mathbb{N}}$$

converge en distribution vers une variable aléatoire Z tel que $Z \sim \mathcal{N}(0,1)$.

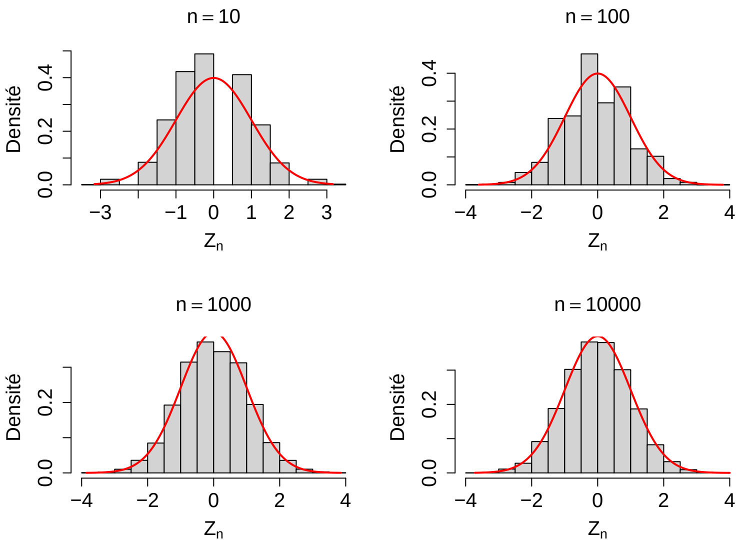


Figure – Illustration du théorème central limite sur base d'un jeu de pile ou face.

Maximum de variables aléatoires i.i.d.

- Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles i.i.d. de fonction de répartition commune F.
- Notons $x_+ = \inf\{x \in \mathbb{R} : F(x) = 1\} = F^{\leftarrow}(1)$.
- Si $M_n = \max(X_1, \ldots, X_n)$, alors $M_n \xrightarrow{D} x_+$. En effet, pour $x \in \mathbb{R}$,

$$\mathbb{P}(M_n \leq x) = \mathbb{P}\left(\bigcap_{i=1}^{n}\{X_i \leq x\}\right) = F(x)^n \to \begin{cases} 0 & \text{si } x < x_+ \\ 1 & \text{si } x \geq x_+. \end{cases}$$

Vers un théorème central limite pour maxima

- Il faut donc également normaliser la suite $(M_n)_{n\in\mathbb{N}}$ si l'on souhaite obtenir une limite en distribution non-dégénérée.
- Ainsi, nous essaierons de trouver les limites possibles non-dégénérées G telles que

$$\mathbb{P}\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \to G(x), \qquad \forall x \in \mathbb{R},$$

où $(a_n)_{n\in\mathbb{N}} > 0$ et $(b_n)_{n\in\mathbb{N}} \in \mathbb{R}$ sont les constantes de normalisation.

Le résultat Fisher et Tippett

En 1928 [Fisher and Tippett, 1928], Ronald A. Fisher (1890–1962) et Leonard H.C. Tippett (1902–1985) proposent une solution au problème.

Théorème de Fisher et Tippett

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires réelles i.i.d. de fonction de répartition commune F. Supposons qu'il existe des constantes $(a_n)_{n\in\mathbb{N}} > 0$ et $(b_n)_{n\in\mathbb{N}} \in \mathbb{R}$ telles que

$$\lim_{n\to\infty} F^n(a_n x + b_n) = G(x),$$

pour tout $x \in \mathbb{R}$ tel que G soit continue en ce point et où G est une fonction de répartition non-dégénérée, alors, G satisfait forcément, pour tout $x \in \mathbb{R}$,

$$G(x) = G_{\mu,\sigma,\gamma}(x) = \exp\left[-\left(1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right)_+^{-1/\gamma}\right],$$

où $\mu \in \mathbb{R}$ est un paramètre de localisation, $\sigma > 0$ est un paramètre d'échelle et $\gamma \in \mathbb{R}$ est l'*indice de valeur extrême*.

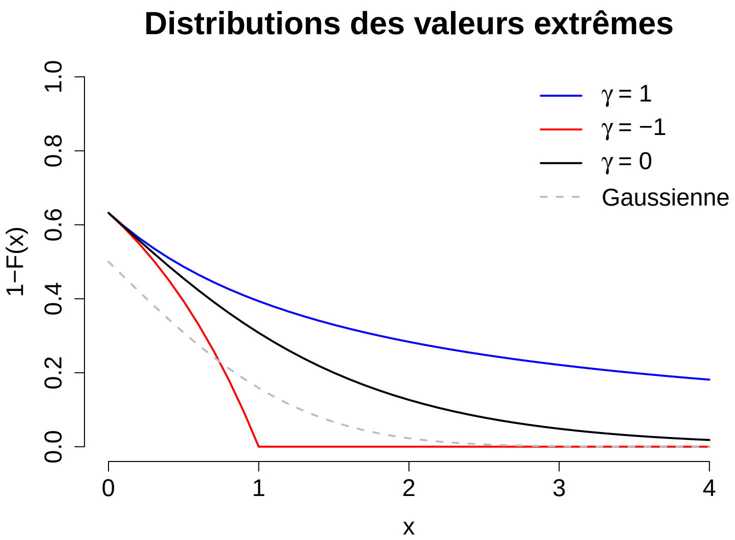


Figure – Fonctions de survie des distributions des valeurs extrêmes en fonction du signe de l'indice ($\mu = 0$ et $\sigma = 1$ dans chaque cas).

Un outil essentiel : le théorème de convergence des types

Notre démonstration, qui suivra essentiellement celle de [Resnick, 1987], utilise le théorème fondamental suivant dont la démonstration n'est pas présentée ici.

Théorème de convergence des types

Soient U et V deux fonctions de répartition non-dégénérées. Soit $(F_n)_{n\in\mathbb{N}}$ une suite de fonctions de répartitions telles que pour tout point de continuité $x \in \mathbb{R}$ de U et $y \in \mathbb{R}$ de V on a

$$\lim_{n\to\infty} F_n(a_n x + b_n) = U(x) \quad \text{et} \quad \lim_{n\to\infty} F_n(\alpha_n y + \beta_n) = V(y),$$

où $(a_n)_{n\in\mathbb{N}} > 0$, $(b_n)_{n\in\mathbb{N}} \in \mathbb{R}$, $(\alpha_n)_{n\in\mathbb{N}} > 0$ et $(\beta_n)_{n\in\mathbb{N}} \in \mathbb{R}$ sont des constantes. Alors,

$$\lim_{n\to\infty} \frac{\alpha_n}{a_n} = A > 0 \quad \text{et} \quad \lim_{n\to\infty} \frac{(\beta_n - b_n)}{a_n} = B \in \mathbb{R},$$

et pour tout $x \in \mathbb{R}$,

$$V(x) = U(Ax + B).$$

Démonstration I

Pour tout $x \in \mathbb{R}$, notons $[x]$ le plus grand nombre entier inférieur ou égal à x.
Soit $t > 0$. On a, pour tout $x \in \mathbb{R}$ en lequel G est continue,

$$\lim_{n\to\infty} F^{[nt]}(a_{[nt]}x + b_{[nt]}) = G(x),$$

mais aussi,

$$\lim_{n\to\infty} F^{[nt]}(a_n x + b_n) = \lim_{n\to\infty} \left(F^n(a_n x + b_n)\right)^{[nt]/n} = \left(G(x)\right)^t .$$

Il suit du théorème de convergence des types que pour tout $t > 0$,

$$\lim_{n\to\infty} \frac{a_{[nt]}}{a_n} = \alpha(t) > 0 \quad \text{et} \quad \lim_{n\to\infty} \frac{b_n - b_{[nt]}}{a_{[nt]}} = \beta(t) \in \mathbb{R},$$

et, de plus, pour tout $x \in \mathbb{R}$,

$$G^t(x) = G(\alpha(t)x + \beta(t)).$$

Démonstration II

Soient $t > 0$ et $s > 0$. Il suit de l'équation précédente que pour tout $x \in \mathbb{R}$,

$$G^{ts}(x) = G(\alpha(ts)x + \beta(ts)),$$

mais on a aussi

$$G^{ts}(x) = (G^s(x))^t = G^t(\alpha(s)x + \beta(s)) = G(\alpha(t)\alpha(s)x + \alpha(t)\beta(s) + \beta(t)).$$

Un argument simple permet de déduire que les fonctions α et β doivent satisfaire les équations suivantes pour tout $t, s > 0$:

$$\begin{cases} \alpha(ts) = \alpha(t)\alpha(s) \\ \beta(ts) = \alpha(t)\beta(s) + \beta(t) = \alpha(s)\beta(t) + \beta(s). \end{cases}$$

Démonstration III

La première équation est une variation simple de *l'équation fonctionnelle de Cauchy* :

$$\boxed{f(x+y) = f(x) + f(y), \qquad \forall x, y \in \mathbb{R}.}$$

En effet, si f est solution de l'équation de Cauchy, alors $\alpha(t) = \exp[f(\log(t))]$ est solution de l'équation précédente.

Démonstration III

La première équation est une variation simple de *l'équation fonctionnelle de Cauchy* :

$$\boxed{f(x+y) = f(x) + f(y), \qquad \forall x, y \in \mathbb{R}.}$$

En effet, si f est solution de l'équation de Cauchy, alors $\alpha(t) = \exp[f(\log(t))]$ est solution de l'équation précédente.
Les uniques solutions *mesurables* de l'équation de Cauchy sont de la forme $f(x) = \theta x$, pour un certain $\theta \in \mathbb{R}$, voir par exemple [Senata, 1976, page 10].
La fonction α est limite de fonctions qui sont constantes par parties, donc mesurables.
On déduit que pour tout $t > 0$,

$$\alpha(t) = \exp(\theta \log(t)) = t^{\theta}, \qquad \theta \in \mathbb{R}.$$

Il faut maintenant distinguer les cas selon le signe de θ. On ne considérera ici que le cas $\theta < 0$ (le cas $\theta > 0$ est tout à fait similaire et le cas $\theta = 0$ est plus simple).

Démonstration IV

Posons $\theta = -\gamma$, où $\gamma > 0$.
Il suit des développements précédents que la fonction β doit satisfaire, pour tout $t, s > 0$,

$$t^{-\gamma}\beta(s) + \beta(t) = s^{-\gamma}\beta(t) + \beta(s),$$

ce qui se réécrit, pour tout $t \neq 1$ et $s \neq 1$,

$$\frac{\beta(t)}{1 - t^{-\gamma}} = \frac{\beta(s)}{1 - s^{-\gamma}}.$$

Dès lors, la fonction $t \in (0, \infty) \setminus \{1\} \mapsto \beta(t)/(1 - t^{-\gamma})$ est constante, notons $C \in \mathbb{R}$ sa valeur, et l'on déduit que

$$\beta(t) = C(1 - t^{-\gamma}), \qquad t \in (0, \infty) \setminus \{1\}.$$

Démonstration V

La fonction de répartition G doit alors satisfaire, pour tout $x \in \mathbb{R}$ et $t \in (0, \infty) \setminus \{1\}$,

$$G^t(x) = G(\alpha(t)x + \beta(t)) = G\left(t^{-\gamma}(x - C) + C\right).$$

La fonction $H : x \in \mathbb{R} \mapsto G(x - C) \in [0, 1]$ satisfait alors

$$H^t(x) = H(t^{-\gamma}x), \qquad \forall x \in \mathbb{R}, t > 0.$$

Notons que cela implique $H(0) \in \{0, 1\}$. Le cas $H(0) = 1$ est rapidement exclu par hypothèse que G (et donc H) soit non-dégénérée.
Ainsi, $H \equiv 0$ sur $(-\infty, 0]$ et il en va de même pour G sur $(-\infty, -C]$.

Démonstration VI

Notre équation sur H montre également que $H(1) \in (0,1)$. En choisissant $x = 1$ dans celle-ci, on trouve alors

$$H(1)^t = H(t^{-\gamma}), \qquad \forall t > 0.$$

On pose alors $t^{-\gamma} = u$ et on obtient que pour tout $u > 0$,

$$H(u) = H(1)^{u^{-1/\gamma}} = \exp\left(-u^{-1/\gamma}\left[-\log(H(1))\right]\right).$$

Soit $x > -C$. En notant $\sigma = \gamma(-\log(H(1)))^{\gamma} > 0$ et $\mu = \sigma/\gamma - C \in \mathbb{R}$, on trouve

$$\begin{aligned} G(x) = H(x+C) &= \exp\left(-(x+C)^{-1/\gamma}\left[-\log(H(1))\right]\right) \\ &= \exp\left[-\left(1+\gamma\left(\frac{x-\mu}{\sigma}\right)\right)^{-1/\gamma}\right]. \end{aligned}$$

Cela conclut la démonstration. $\square$

La théorie des valeurs extrêmes en pratique

- La ville nous demande de construire des digues suffisamment hautes de telle sorte à ce que la hauteur annuelle maximale de l'eau soit inférieure à la hauteur de ces digues dans 99.99% des cas.

- Il est impossible de se reposer sur une estimation empirique dans ce cas car on dispose uniquement de 65 données.
- A la place, on propose de se baser sur la théorie des valeurs extrêmes et l'on va supposer que nos données $\{M_i : i = 1, \ldots, 65\}$ suivent approximativement une distribution des valeurs extrêmes $G = G_{\mu,\sigma,\gamma}$.

Estimation des paramètres

- L'estimation des paramètres est faite par maximum de vraisemblance, voir [Beirlant et al., 2004, Section 5.1.2], via la fonction `fgev` du package R evd.

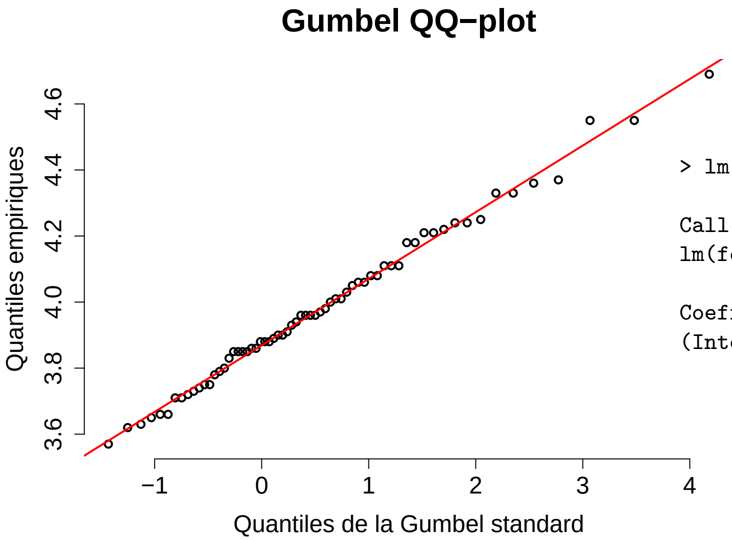
```
> fgev(portpirie)

Call: fgev(x = portpirie)
Deviance: -8.678117

Estimates
     loc     scale     shape
 3.87475   0.19805  -0.05012

Standard Errors
    loc    scale    shape
0.02793  0.02025  0.09826

Optimization Information
  Convergence: successful
  Function Evaluations: 33
  Gradient Evaluations: 8
```

Quel modèle choisir ?

- L'estimation $\hat{\gamma} = -0.05012$ du paramètre de forme est très proche de zéro.
- On préfère être prudent et vérifier si les données collent au modèle $\{\Lambda_{\mu,\sigma}(x) = G_{\mu,\sigma,\gamma=0}(x) = \exp(-\exp(-(x-\mu)/\sigma)), x \in \mathbb{R}\}$, appelé *modèle de Gumbel*.
- Pour cela, une technique classique est d'avoir recours au *QQ-plot*. L'idée derrière ce graphe est que le quantile x_p d'ordre $p \in (0,1)$ de $\Lambda_{\mu,\sigma}$, qui est solution de $\Lambda_{\mu,\sigma}(x_p) = p$ satisfait

$$x_p = \mu + \sigma\left[-\log(-\log(p))\right].$$

  Autrement dit, le quantile d'ordre p d'une distribution de Gumbel $\Lambda_{\mu,\sigma}$ est en relation affine avec celui d'une distribution de Gumbel $\Lambda_{\mu=0,\sigma=1}$ dite *standard*.
- Ainsi, si l'on réalise un graphique des quantiles empiriques de nos données en fonction des quantiles de la distribution $\Lambda_{\mu=0,\sigma=1}$ pour plusieurs valeurs de $p \in (0,1)$, une droite est supposée apparaitre.




```
> lm(q_emp~q_th)

Call:
lm(formula = q_emp ~ q_th)

Coefficients:
(Intercept)         q_th
     3.8690       0.2016
```

Hauteur de la digue

- Le modèle semblant relativement bien représenter nos données, le calcul de la hauteur h de la digue peut être effectué.
- h doit satisfaire
$$\Lambda_{\hat{\mu},\hat{\sigma}}(h) = 0.9999,$$
ou, dit autrement,
$$\begin{aligned} h = x_{0.9999} &= \hat{\mu} + \hat{\sigma}\,[-\log(-\log(0.9999))] \\ &= 3.87475 + 0.19805\,[-\log(-\log(0.9999))] \\ &= 5.698848. \end{aligned}$$
- Notons que $h > 4.69$, la hauteur que l'eau avait atteinte le jour des inondations de Port Pirie en 1934.

Domaine d'attraction

- Dans notre théorème centrale limite pour maxima, nous avons caractérisé les limites possibles $G = G_\gamma$ de maxima (normalisés) de suites de variables aléatoires i.i.d.
- Deux questions demeurent encore :
  1. Étant donné G_γ, pour quelles fonctions de répartition F existe-t-il des constantes $(a_n)_{n\in\mathbb{N}} > 0$ et $(b_n)_{n\in\mathbb{N}} \in \mathbb{R}$ telles que
     $$\lim_{n\to\infty} F^n(a_n x + b_n) = G_\gamma(x)$$
     pour tout $x \in \mathbb{R}$? Si une telle fonction F existe, on dira que F est *dans le domaine d'attraction* de G_γ.
  2. Comment caractériser les constantes $(a_n)_{n\in\mathbb{N}} > 0$ et $(b_n)_{n\in\mathbb{N}} \in \mathbb{R}$?
- Nous verrons les réponses à ces questions dans le cas $\gamma > 0$, un problème qui avait déjà étudié par Boris V. Gnedenko (1912–1995) dans son célèbre papier (en français !) [Gnedenko, 1943]. Les outils utilisés constituent un cas particulier de la théorie générale de *variation régulière* développée plus largement dans la thèse de Laurens de Haan (né en 1937) [de Haan, 1970].

Variation régulière d'une fonction d'une variable réelle I

Variation régulière [Bingham et al., 1987, Chapitre 1]

Soit $f : (0, \infty) \to (0, \infty)$ mesurable. On dit que f est à variation régulière (à l'infini) si pour tout $x > 0$, la limite suivante existe :

$$\lim_{t \to \infty} \frac{f(tx)}{f(t)} = g(x).$$

Dans ce cas, on a forcément que $g(x) = x^\rho$ pour un certain $\rho \in \mathbb{R}$ et on dit que f est à variation régulière d'indice ρ, ce qui sera noté $f \in \mathrm{RV}_\rho$. Si $\rho = 0$, on dira que f est à variation lente.

Variation régulière d'une fonction d'une variable réelle II

- On remarque que si $f \in \mathrm{RV}_\rho$, alors forcément pour tout $x > 0$, on a

$$f(x) = x^\rho L(x)$$

  où L est une fonction à variation lente.
- Dit autrement, une fonction $f \in \mathrm{RV}_\rho$ se comporte asymptotiquement comme une fonction puissance multipliée par un facteur qui varie *plus lentement* qu'une fonction puissance.
- Quelques exemples et contre-exemples :

Fonctions (sur $(0, \infty)$)	Variation
$f(x) = x^\rho$	Variation régulière d'indice ρ
f satisfait $\lim_{x \to \infty} f(x) = C > 0$	Variation lente
$f(x) = \log(x)$	Variation lente
$f(x) = x^\rho \log(x)$	Variation régulière d'indice ρ
$f(x) = \exp(x)$	Pas de variation régulière
$f(x) = 1 - \exp(-x^{-\alpha})$ où $\alpha > 0$	Variation régulière d'indice $-\alpha$

Domaine d'attraction de G_γ avec $\gamma > 0$

[Gnedenko, 1943, Théorème 4]

1. Une fonction de répartition F est dans le domaine d'attraction d'une distribution des valeurs extrêmes G_γ pour $\gamma > 0$ si et seulement si $1 - F \in \mathrm{RV}_{-1/\gamma}$.
2. Si l'une des conditions (équivalentes) du point 1 est satisfaite, alors, pour tout $x \in \mathbb{R}$,

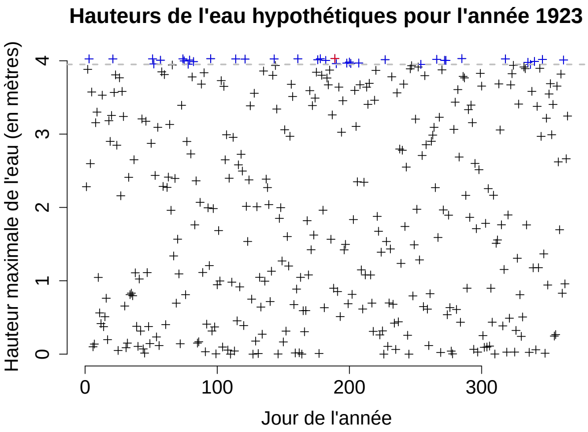
$$\lim_{n\to\infty} F^n(a_n x) = G_{\mu=1,\sigma=\gamma,\gamma}(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \exp(-x^{-1/\gamma}) & \text{si } x > 0 \end{cases},$$

où $a_n = (1/(1-F))^{\leftarrow}(n) = \inf\{x \in \mathbb{R} : 1/(1-F(x)) \geq n\}$.

- Nous n'aborderons pas la démonstration complète mais seulement certaines parties. La preuve entière est à trouver chez [Resnick, 1987, Proposition 1.11].

Application statistique : la méthode Peak Over Thresholds (PoT) I

- Supposons que dans l'analyse des données de Port Pirie, nous ne disposions pas seulement des maxima annuels mais de la hauteur de l'eau maximale quotidienne pour chaque jour entre 1923 et 1987, i.e., $n = 365 * 65 = 23725$ données.
- Par exemple, pour l'année 1923, on pourrait avoir :



Application statistique : la méthode Peak Over Thresholds (PoT) II

- La variation régulière permet d'utiliser toutes les données bleues (soit 36 données, environ 10% des 365 données de l'année) au lieu de se restreindre uniquement à l'information contenue dans le maximum (en rouge).
- Evidemment, cela fonctionne aussi pour les autres années et on obtiendrait finalement $k = 36 * 65 = 2340$ données au lieu des 65 maxima initiaux.
- Par le théorème de Gnedenko, si l'on suppose que M_i (le maximum de l'année $i = 1, \ldots, 65$) est distribué selon une distribution des valeurs extrêmes G_γ avec $\gamma > 0$, alors, pour tout $x > 0$,

$$\lim_{t \to \infty} \frac{1 - F(tx)}{1 - F(t)} = x^{-1/\gamma}.$$

En termes de variables aléatoires, cela donne,

$$\lim_{t \to \infty} \frac{\mathbb{P}(X > tx)}{\mathbb{P}(X > t)} = x^{-1/\gamma}.$$

Application statistique : la méthode Peak Over Thresholds (PoT) III

- En prenant $t = X_{n-k,n}$, on trouve que pour tout $x > 0$,

$$\mathbb{P}(X > tx) \approx P(X > t)x^{-1/\hat{\gamma}} = \frac{k}{n}x^{-1/\hat{\gamma}},$$

  où $\gamma > 0$ peut être estimé par nos k données, via, par exemple, l'*estimateur de Hill* donné par

$$\hat{\gamma} = \frac{1}{k}\sum_{i=1}^{k}\log\left(\frac{X_{n-i+1,n}}{X_{n-k,n}}\right).$$

  Voir [Beirlant et al., 2004, Chapitre 4] pour plus de détails.
- Il est possible de généraliser ces arguments au cas $\gamma \in \mathbb{R}$ mais cela requiert des notions plus générales que la variation régulière [Resnick, 1987, Section 0.4.3].

Éléments de démonstration du théorème de Gnedenko I

On va montrer qu'il existe des constantes $(a_n)_{n\in\mathbb{N}} > 0$ telles que

$$\lim_{n\to\infty} F^n(a_n x) = \begin{cases} 0 & \text{si } x \leq 0 \\ \exp(-x^{-1/\gamma}) & \text{si } x > 0 \end{cases},$$

si et seulement $1 - F \in \mathrm{RV}_{-1/\gamma}$ et $a_n \sim (1/(1-F))^{\leftarrow}(n)$ quand $n \to \infty$.

$\boxed{\Longrightarrow}$ Commencer par montrer $x_+ = F^{\leftarrow}(1) = +\infty$. En effet, sinon, on doit avoir $a_n x \to x_+$ quand $n \to \infty$ pour n'importe quel $x > 0$, ce qui impliquerait $a_n \to 0$ et donc $x_+ = 0$, ce qui est impossible au vue de la limite ci-dessus. Ainsi $x_+ = +\infty$.
De plus, on a forcément $a_n \to \infty$ quand $n \to \infty$ sans quoi, quitte à passer à une sous-suite $(a_{n_k})_{k\in\mathbb{N}}$, on aurait $a_{n_k} \leq K$ pour un certain $K > 0$ et pour tout $k \in \mathbb{N}$. Cela impliquerait

$$0 < \exp(-1) = \lim_{k\to\infty} F^{n_k}(a_{n_k}) \leq \lim_{k\to\infty} F^{n_k}(K) = 0,$$

qui est une contradiction.

Éléments de démonstration du théorème de Gnedenko II

Prenant le logarithme de la limite de base, on a, pour tout $x > 0$

$$\lim_{n\to\infty} n \log F(a_n x) = -x^{-1/\gamma}.$$

Notre raisonnement précédent montre que $a_n x \to \infty$ pour tout $x > 0$ quand $n \to \infty$ donc

$$\lim_{n\to\infty} \frac{-\log F(a_n x)}{1 - F(a_n x)} = 1$$

car $-\log(z) \sim 1 - z$ quand $z \nearrow 1$. Donc, on doit avoir pour tout $x > 0$,

$$\lim_{n\to\infty} n\,[1 - F(a_n x)] = x^{-1/\gamma}.$$

Nous allons montrer que cela suffit à démontrer que $1 - F \in \mathrm{RV}_{-1/\gamma}$.

Éléments de démonstration du théorème de Gnedenko III

Soit $t > 0$. Alors, comme $a_n \to \infty$ quand $n \to \infty$, on peut définir

$$n(t) = \inf\{m \in \mathbb{N} : a_{m+1} > t\}$$

qui satisfait

$$a_{n(t)} \leq t < a_{n(t)+1}.$$

Il suit que pour tout $x > 0$,

$$\left(\frac{n(t)+1}{n(t)}\right)\left(\frac{n(t)[1-F(a_{n(t)}x)]}{(n(t)+1)[1-F(a_{n(t)+1})]}\right) \leq \frac{1-F(tx)}{1-F(t)} \leq \left(\frac{n(t)}{n(t)+1}\right)\left(\frac{(n(t)+1)[1-F(a_{n(t)+1}x)]}{n(t)[1-F(a_{n(t)})]}\right)$$

En prenant la limite $t \to \infty$, on obtient que pour tout $x > 0$,

$$\lim_{t\to\infty} \frac{1-F(tx)}{1-F(t)} = x^{-1/\gamma}.$$

Éléments de démonstration du théorème de Gnedenko IV

Il reste à montrer que $a_n \sim (1/(1-F))^{\leftarrow}(n)$ quand $n \to \infty$. Cela suit du fait que pour tout $x > 0$,

$$\lim_{n\to\infty} n^{-1}\left[\frac{1}{1-F(a_n x)}\right] = x^{1/\gamma},$$

ce qui, par le caractère croissant des fonctions présentes dans cette limite, implique la convergence ponctuelle de leurs « inverses », c'est-à-dire, pour tout $x > 0$,

$$\lim_{n\to\infty} \frac{(1/(1-F))^{\leftarrow}(nx)}{a_n} = x^{\gamma}.$$

Choisir $x = 1$ permet de conclure la première partie de la démonstration.

Éléments de démonstration du théorème de Gnedenko V

$\boxed{\Longleftarrow}$ Si l'on suppose maintenant $1 - F \in \mathrm{RV}_{-1/\gamma}$ et $a_n \sim (1/(1-F))^{\leftarrow}(n)$, alors pour tout $x > 0$,

$$\lim_{n\to\infty} \frac{1 - F(a_n x)}{1 - F(a_n)} = x^{-1/\gamma}.$$

On peut facilement montrer que $1 - F(a_n) \sim n^{-1}$ quand $n \to \infty$, donc, on a pour tout $x > 0$,

$$\lim_{n\to\infty} n\left[1 - F(a_n x)\right] = x^{-1/\gamma}$$

qui correspond à ce qu'on avait obtenu plus tôt en transformant la limite initiale. Il suffit alors d'effectuer les transformations inverses pour conclure cette partie de la démonstration. $\square$

Conclusions et perspectives

- On a montré que la théorie des valeurs extrêmes fournit un cadre théorique intéressant pour répondre à des questions liées à des évènements pour lesquels peu ou pas de données historiques sont disponibles, et ce, dans un cadre non-paramétrique via la théorie asymptotique.
- Beaucoup de points n'ont bien sûr pas été abordés :
  - qualité de l'approximation asymptotique
  - théorie statistique des estimateurs des paramètres des distributions des valeurs extrêmes
  - variation régulière dans le cas général $\gamma \in \mathbb{R}$
  - données non i.i.d.
  - données multivariées
  - données fonctionnelles
  - ...
- Ces points (et beaucoup d'autres) font encore l'objet de nombreux articles actuels et sont un sujet de recherche très actif, autant du point de vue théorique que du point de vue appliqué !

Références I

Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004).
Statistics of Extremes : Theory and Applications.
Wiley Series in Probability and Statistics. Wiley.

Bingham, N. H., Goldie, C. M., and Teugels, J. (1987).
Regular Variation, volume 27 of *Encyclopedia of Mathematics and its Applications*.
Cambridge University Press.

de Haan, L. (1970).
On Regular Variation and Its Application to the Weak Convergence of Sample Extremes, volume 32 of *Mathematical Centre tracts*.
Mathematisch Centrum.

Références II

Fisher, R. A. and Tippett, L. (1928).
Limiting forms of the frequency distribution of the largest or smallest members of a sample.
Proceedings of the Cambridge Philosophical Society, 24 :180–190.

Gnedenko, B. V. (1943).
Sur la distribution limite du terme maximum d'une serie aleatoire.
The Annals of Mathematics, 44(3) :423–453.

Resnick, S. I. (1987).
Extreme values, regular variation and point processes.
Springer Series in Operations Research and Financial Engineering. Springer New York, NY.

Senata, E. (1976).
Regularly Varying Functions.
Lecture Notes in Mathematics. Springer Berlin, Heidelberg.