

Statistiques des événements extrêmes et application

Jean-Noel Bacro

Mathieu Ribatet

Résumé

L'objectif de ce projet est de mettre en pratique la théorie présentée en cours. Vous essaieriez donc de modéliser le comportement extrêmes des températures en Suisse à l'aide de la théorie des valeurs extrêmes. Votre analyse sera univariée puis bivariée. Comme précisé en cours, vous pouvez travailler en binôme mais pas plus. La rédaction du projet devra faire apparaître votre raisonnement et la démarche utilisée. Une attention particulière sera portée à la qualité de la rédaction et de la présentation des résultats (graphiques, tableaux, ...). Le code des programmes sera fourni en annexe.

1 Les données

Les données que vous devez analyser correspondent aux températures journalières enregistrées en 20 stations météorologiques. Afin de faciliter votre prise en main de ce projet, nous avons regroupé toutes les données nécessaires au sein d'un unique fichier `swissTemperature.RData` que vous importerez dans R à l'aide de la commande

```
load("swissTemperature.RData")
```

Ceci aura pour but d'importer 3 objets R

```
> ls()
[1] "swiss"          "swissTemp"      "swissTempCoord"
```

L'objet `swissTempCoord` est une matrice contenant les coordonnées géographiques des stations météorologiques dont nous disposons.

```
> swissTempCoord
      Lon (km) Lat (km) Alt (km)
Arosa      770.730 183.320   1.840
Bad Ragaz   756.900 209.370   0.496
Basel-Binningen 610.850 265.620   0.316
Bern-Liebefeld 598.610 197.470   0.565
Chateau d'Oex 577.200 147.310   0.985
Davos-Dorf  783.580 187.480   1.590
Engelberg   674.150 186.060   1.035
Gd-St-Bernard 579.200  79.720   2.472
Jungfrauoch 641.930 155.275   3.580
Locarno-Monti 704.160 114.350   0.366
Lugano      717.880  95.870   0.273
Montana     603.600 129.160   1.508
Montreux-Clarens 558.560 143.600   0.405
Neuchatel   563.150 205.600   0.485
Oeschberg-Koppigen 613.250 219.525   0.483
Santis      744.100 234.900   2.490
Zurich-MeteoSchweiz 685.125 248.090   0.556
```



FIGURE 1 – Positions géographiques des 20 stations météorologiques en Suisse.

L'objet `swissTemp` est une matrice contenant les températures journalières observées en chacune des stations précédentes de 1961 à 2006. Attention toutefois aux valeurs manquantes...

	year	month	day	Arosa	Bad Ragaz	Basel-Binningen	Bern-Liebefeld	...
1	1961	1	1	-3.0	5.0	2.1	1.7	
2	1961	1	2	-4.0	5.8	7.5	2.0	
3	1961	1	3	0.5	8.3	12.8	3.3	
4	1961	1	4	-4.4	3.2	6.8	3.1	
5	1961	1	5	-6.0	0.6	5.3	2.6	
.								
.								
.								
16583	2006	5	27	12.7	21.5	22.8	22.2	
16584	2006	5	28	10.7	19.1	19.3	21.5	
16585	2006	5	29	10.7	17.2	14.2	16.6	
16586	2006	5	30	-0.4	NA	11.9	10.4	
16587	2006	5	31	0.8	11.0	12.2	9.8	
16588	2006	6	1	2.7	13.3	10.3	11.9	

Enfin, l'objet `swiss` est en fait une fonction permettant de représenter la Suisse. Nous pouvons donc nous servir de cette fonction afin de représenter graphiquement les stations météorologiques dont nous disposons pour ce projet. Par exemple le code suivant¹ nous a permis de générer la Figure 1.

```
par(mar = rep(0, 4))
swiss(asp = 1)
text(swissTempCoord, rownames(swissTempCoord))
```

1. L'option `|asp = 1|` n'est pas nécessaire mais permet de ne pas déformer la Suisse lors du redimensionnement de la fenêtre graphique.

2 Objectifs

1. On considère Neuchatel comme une station de référence. Procédez à une analyse univariée des extrêmes de cette station. Les approches GEV et GPD pourront être utilisées... et discutées!
2. On considère les stations distantes de moins de 60 km de Neuchatel. En prenant l'une de ces stations, vous poursuivrez votre étude en faisant une analyse bivariée des deux stations. On pourra notamment s'intéresser à la prédiction de quantiles conditionnels du type

$$P(X \leq x \mid Y > y)$$

avec $y = 35^\circ\text{C}$ et où Y représente la température à Neuchatel et X celle de la station choisie.

3. Reprenez la question précédente en choisissant maintenant une station à plus de 200 km de Neuchatel.
4. Comparez et discutez les résultats obtenus aux deux questions précédentes.
5. *Question bonus* : Si le temps vous le permet vous pourrez vous initier à l'analyse spatiale des extrêmes en supposant un modèle très simple

$$\{Y(x_1), \dots, Y(x_{20})\} \sim \prod_{i=1}^{20} \text{GEV}\{\mu(x_i), \sigma(x_i), \xi(x_i)\},$$

où les paramètres de la GEV seront représentés par des surfaces de réponses comme par exemple

$$\mu(x) = \beta_0 + \beta_1 \text{lon}(x), \quad x \in \mathcal{X},$$

et $\text{lon}(x)$ représente la longitude au point x .

3 Rédaction

Pour ce projet vous serez évalués sur votre rigueur scientifique mais aussi sur la rédaction. Pour ce dernier point vous porterez donc une attention particulière à la qualité des figures, tableaux ainsi qu'à la structure de votre document.

Vous trouverez des conseils utiles sur la rédaction scientifique à partir de ce lien.

Bon courage et travail!