

Chapitre 2

Tests en amont et tests en aval d'une analyse fréquentielle

I. Tests en amont d'une analyse fréquentielle (tests d'hypothèse)

I.1. Introduction

L'analyse fréquentielle des données hydro climatologiques (pluie, débit ou autre) nécessite que les données au niveau des stations de mesure soient indépendantes (sans autocorrélations ni tendances) et qu'elles soient identiquement distribuées.

L'indépendance se testera en utilisant les coefficients d'autocorrélation du lag-1 au lag-5. La tendance sera vérifiée à l'aide du test non paramétrique Mann-Kendall (MK) (Mann, 1945 ; Kendall, 1975) tel qu'il est recommandé par l'Organisation Météorologique Mondiale (WMO, 1988). Une autre exigence est que les séries temporelles aux différentes stations soient homogènes, cette homogénéité des séries sera réalisée par les tests non paramétriques de Wilcoxon-Mann-Whitney (Wilcoxon, 1945 ; Mann and Whitney, 1947) en outre les dates de rupture probables seront détectées par le test de Pettitt (Pettitt, 1979). Pour se rassurer que les valeurs ponctuelles dans chaque série temporelle soient aléatoires, le test de Wald-Wolfowitz (Wald and Wolfowitz, 1940) a été adopté.

NB : Un niveau de signification de $\alpha = 0,05$ a été utilisé dans tous les tests de prétraitement des données entrepris dans cette section d'étude.

I.2. Test d'homogénéité de Wilcoxon-Mann-Whitney

Le test de **Wilcoxon-Mann-Whitney** (ou test U de **Mann-Whitney** ou encore test de la somme des rangs de **Wilcoxon**) est un test statistique non paramétrique qui permet de tester l'hypothèse selon laquelle la distribution des données est la même dans deux groupes.

Il a été proposé par **Frank Wilcoxon** en 1945 (Wilcoxon, 1945) et par **Henry Mann** et **Donald Ransom Whitney** en 1947 (Mann et Whitney, 1947).

L'énorme avantage de ce test est sa simplicité, même si de ce fait son utilisation est limitée. Comme tous les tests statistiques, il consiste à partir de ce qui est observé à mettre en évidence un événement dont on connaît la loi de probabilité (au moins sa forme asymptotique). La valeur obtenue, si elle est peu probable selon cette loi, suggèrera de rejeter l'hypothèse nulle.

Le test U de **Mann-Whitney** calcule une p-value qui se base non pas sur les données brutes, mais sur leurs rangs. Les rangs en question correspondent à l'emplacement de chaque donnée au sein de la série lorsque cette série est ordonnée par ordre croissant.

Les distributions des rangs sont beaucoup moins affectées par l'existence de valeurs extrêmes que ne le sont les distributions des valeurs brutes.

Le test de Mann-Whitney peut se baser sur une métrique **W** (en lien avec la dénomination « test de **Wilcoxon** ») qui correspond à :

$$W = W_A = \sum_{i=1}^n 1_A R_i - n_A(n_A + 1) / 2 \quad (1)$$

On peut aussi définir W_B :

$$W = W_B = \sum_{i=1}^n 1_B R_i - n_B(n_B + 1) / 2 \quad (2)$$

La notation $\sum_{i=1}^n 1_A R_i$ correspond simplement à la somme des rangs des individus du groupe A.

La notation $\sum_{i=1}^n 1_B R_i$ correspond simplement à la somme des rangs des individus du groupe B.

Les nombres $n_A(n_A + 1) / 2$ et $n_B(n_B + 1) / 2$ correspondent en fait à ce que serait la somme des rangs si tous les individus du groupe A (respectivement B) avaient les rangs $1, 2, 3, \dots, n_A$ (respectivement $1, 2, 3, \dots, n_B$).

W est ainsi d'autant plus grand que les individus du groupe A présentent des rangs élevés au sein de l'ensemble des individus.

Alternativement, on pourrait considérer la métrique U (en lien avec la dénomination « test U de Mann-Whitney ») qui correspond à :

$$U = \min(W_A, W_B) \quad (3)$$

Cette métrique est quant à elle d'autant plus faible que l'un ou l'autre des groupes présente des rangs bas.

En l'occurrence, par construction on a toujours :

$$W_A + W_B = n_A n_B \quad (4)$$

La distribution empirique de W nous permet de calculer la **p-value** correspondant au test de l'hypothèse si les deux groupes A et B proviennent ou non de la même population, c'est-à-dire vérifier au seuil de 5% si l'échantillon (la série) formé par A et B est homogène ou non.

Si la **p-value** est supérieure à **0,05**, la série est homogène et vice versa.

Exemple : Pluie annuelle à la station de Batna ($N = 69$, de 1936 à 2004).

$W = 532$, $p\text{-value} = 0,4556 > 0,05$ la série est homogène.

1.3. Test de rupture de Pettitt

Le test de **Pettitt** (Pettitt, 1979), est un test non paramétrique qui a été utilisé dans un certain nombre d'études hydro-climatologiques pour détecter des changements brusques dans la moyenne de la distribution de la variable d'intérêt. Ce test est basé sur le test à deux échantillons de Mann-Whitney (test basé sur le rang) et permet la détection d'un seul changement à un moment indéterminé. Ce test est souvent utilisé pour détecter des changements dans les extrêmes en raison du manque d'hypothèses de répartition.

Les rangs r_1, \dots, r_n des individus x_1, \dots, x_n sont utilisés dans la statistique :

$$U_k = 2 \sum_{i=1}^k r_i - k(n+1) ; k = 1, \dots, n \quad (5)$$

La statistique du test est le maximum de la valeur absolue du vecteur :

$$U^* = \max |U_k| \quad (6)$$

Une **p-value** correspondante est calculée, si elle est supérieure à **0,05**, la série est homogène et vice versa.

Exemple : Pluie annuelle à la station de Batna ($N = 69$, de 1936 à 2004).

$U^* = 380$, $p\text{-value} = 0,1486 > 0,05$ la série est homogène. Si non la rupture est en 1956.

1.4. Test du caractère aléatoire de Wald-Wolfowitz

Le test de **Wald-Wolfowitz** (Wald et Wolfowitz, 1940), nommé d'après **Abraham Wald** et **Jacob Wolfowitz**, tous deux des statisticiens connus. C'est un test statistique non paramétrique qui vérifie l'hypothèse aléatoire ou non aléatoire pour une séquence de données (séries chronologiques par exemple). Plus précisément, il peut être utilisé pour tester l'hypothèse que les éléments de la séquence sont mutuellement indépendants.

Soit x_1, x_2, \dots, x_n dénote les données échantillonnées, puis la statistique du test de **Wald-Wolfowitz** est calculée comme suit:

$$R = \sum_{i=1}^{n-1} x_i x_{i+1} + x_1 x_n \quad (7)$$

La valeur attendue de R est :

$$E(R) = \frac{s_1^2 - s_2}{n-1} \quad (8)$$

La variance attendue est :

$$V(R) = \frac{s_2^2 - s_4}{n-1} - E(R)^2 + \frac{s_1^4 - 4s_1^2 s_2 + 4s_1 s_3 + s_2^2 - 2s_4}{(n-1)(n-2)} \quad (9)$$

Avec :

$$s_t = \sum_{i=1}^n x_i^t, t = 1, 2, 3, 4 \quad (10)$$

Pour $n > 10$, la statistique du test est normalement distribuée, avec :

$$z = \frac{R - E(R)}{\sqrt{V(R)}} \quad (11)$$

Si la statistique est inférieure à **1.96**, les séries sont aléatoires et stationnaires au seuil de 5%.

Exemple : Pluie annuelle à la station de Batna (N = 69, de 1936 à 2004).

Statistic (z) = -1.7106 < 1,96 la série est aléatoire.

1.5 Test de tendance (stationnarité) de Mann-Kendall

Le test de Mann-Kendall sert à déterminer avec un test non paramétrique si une tendance est identifiable dans une série temporelle qui comprend éventuellement une composante saisonnière. Ce test de tendance non paramétrique est le résultat d'une amélioration du test d'abord été étudié par [Mann \(1947\)](#) puis repris par [Kendall \(1975\)](#). L'hypothèse nulle H_0 de ces tests est qu'il n'y a pas de tendance. Les trois hypothèses alternatives de tendance négative, non nulle ou positive peuvent être choisies. Les tests de Mann-Kendall s'appuient sur le calcul du « τ » de Kendall mesurant l'association entre deux échantillons et lui-même basé sur les rangs à l'intérieur des échantillons.

L'hypothèse nulle est que les données proviennent d'une population avec des réalisations indépendantes et sont distribuées de façon identique. Pour le test bilatéral, l'hypothèse alternative est que les données suivent une tendance monotone. La statistique du test de Mann-Kendall est calculée selon la relation suivante :

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{sgn}(x_j - x_k) \quad (12)$$

Avec « sgn » est la fonction « signe ».

La fonction « signe », ou « signum » en latin, souvent représentée par « sgn » dans les expressions, est une fonction mathématique qui extrait le signe d'un nombre réel, c'est-à-dire que l'image d'un nombre par cette application est 1 si le nombre est strictement positif, 0 si le nombre est nul, et -1 si le nombre est strictement négatif.

La moyenne de S est $\mu = 0$. La variance incluant le terme de correction pour les liens est :

$$\sigma^2 = \left\{ n(n-1)(2n+5) - \sum_{j=1}^p t_j(t_j-1)(2t_j+5) \right\} / 18 \quad (13)$$

Où p est le nombre de groupes liés dans l'ensemble de données et t_j est le nombre de points de données dans le $j^{\text{ème}}$ groupe lié. La statistique S est approximativement distribuée normalement, avec :

$$z = S / \sigma \quad (14)$$

S'il y a continuité, une correction sera utilisée :

$$z = \text{sgn}(S) (|S| - 1) / \sigma \quad (15)$$

La statistique S est étroitement liée à celle de Kendall désignée par τ :

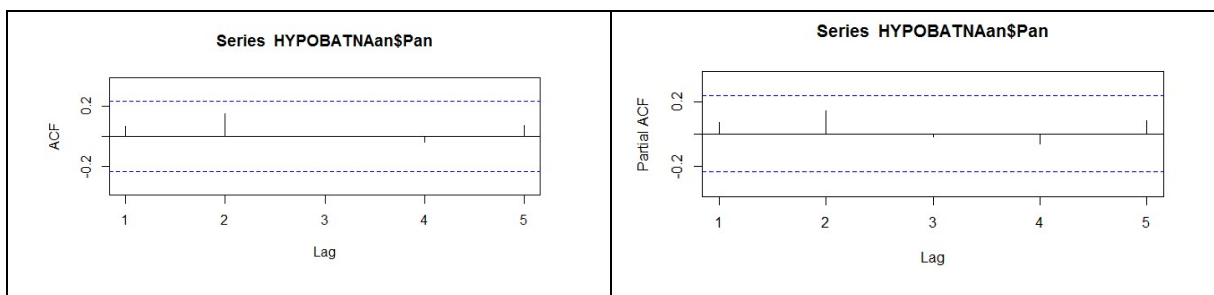
$$\tau = S / D \quad (16)$$

Où

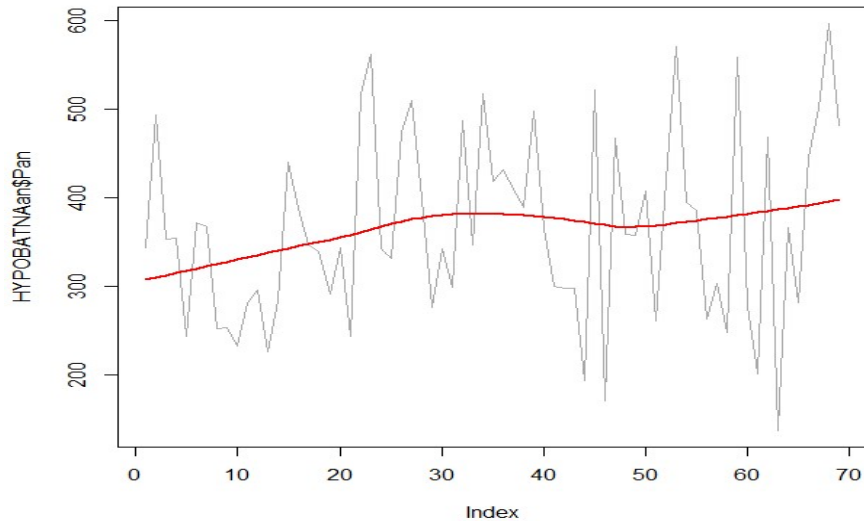
$$D = \left[\frac{1}{2} n(n-1) - \frac{1}{2} \sum_{j=1}^p t_j(t_j-1) \right]^{1/2} \left[\frac{1}{2} n(n-1) \right]^{1/2} \quad (17)$$

Une **p-value** correspondante est calculée, si elle est supérieure à **0,05**, la série ne présente plus de tendance et vice versa.

Exemple : Pluie annuelle à la station de Batna (N=69, de 1936 à 2004).



Selon les deux graphiques (ACF et PACF), il n'existe aucune autocorrelation.



A vu d'oeil, on ne peut pas conclure à propos de la stationnarité (l'existence ou pas d'une tendance)!!

Il faut faire le test statistique non paramétrique de Mann-Kendall.

Tau = 0.1108269, p-value = 0,17976 > 0,05 la série est stationnaire (pas de tendance significative).

II. Tests en aval d'une analyse fréquentielle (tests d'adéquation)

II.1. Introduction

Les lois de probabilité (Normale, lognormale, ou des valeurs extrême (Gumbel, ...) ne seront en mesure d'exploitation que lorsque les tests dits en aval soient réalisés avec succès; on dira donc que les lois ajustées sont adéquates et les tests seront només tests d'adéquation.

Parmi les tests les plus connus, on citera les tests de: Chi-deux de Pearson, Kolmogorov-Smirnov et *Anderson-Darling*. On se limitera dans ce chapitre au dernier test qui sera utilisé dans le logiciel Hydrolab ([Laborde 1998](#)).

II.2. Test d'Anderson-Darling

Le test d'Anderson-Darling ([Stephens, 1974](#)) est utilisé pour déterminer si un échantillon de données provient d'une population (par exemple, une serie hydropluviométrique) dont la répartition est spécifique (par exemple une loi normale de Gauss). Il s'agit d'un test qui donne plus de poids aux queues.

Le test d'Anderson-Darling utilise la distribution spécifique pour calculer les valeurs critiques. Ceci a l'avantage de permettre un test plus sensible et l'inconvénient que les valeurs critiques doivent être calculées pour chaque distribution. Actuellement, des tableaux de valeurs critiques sont disponibles pour les valeurs normales, uniformes, lognormales et exponentielles.

Le fonctionnement du test est selon les étapes qui suivent:

Etape 1: Les hypothèses

Le test d'Anderson-Darling est uniquement un test d'adéquation unilatéral. On aura donc simplement:

- **H0:** Les données suivent la loi ajustée.
- **H1:** Les données ne suivent pas la loi ajustée.

Etape 2: Calculer la valeur pratique

Tout d'abord, on commence par trier nos données de taille N dans l'ordre croissant avec le rang est i . On en déduit la moyenne et l'écart type.

1. On calcule ensuite les données centrées réduites selon la loi ajustée.
2. On utilise ensuite la fonction de répartition de la ajustée centrée réduite pour obtenir les fréquences.
3. On calcule le logarithme népérien de cette valeur.
4. De la même manière on forme F_{n-i+1} puis on en déduit le logarithme népérien.
5. On calcule ensuite la somme S :

$$S = \sum_{i=1}^N \frac{(2i-1)}{N} [\ln F(Y_i) + \ln (1 - F(Y_{N+1-i}))] \quad (18)$$

6. On en déduit la statistique A^2 par le calcul suivant:

$$A^2 = -N \cdot S \quad (19)$$

Etape 3: Calculer la valeur critique

La valeur critique du test a été tabulée et qui dépend de la loi d'ajustement en question, pour le cas particulier d'un ajustement à la loi Normale, les valeurs sont les suivantes:

Niveau de risque	1%	2,5%	5%	10%
Valeur critique	1,035	0,873	0,752	0,631

Etape 4: Calculer la p-value

La p-value est calculée à partir de la Statistique A^2 par interpolation à partir d'une table décrite par [Stephens \(1986\)](#). On utilise le tableau suivant:

Valeur de A^2	p-Value
$A^2 < 0,2$	$1 - e^{(-13,436 + 101,14 * A^2 - 223,73 * A^2^2)}$ (20)
$0,2 \leq A^2 \leq 0,34$	$1 - e^{(-8,318 + 42,796 * A^2 - 59,938 * A^2^2)}$ (21)
$0,34 \leq A^2 < 0,6$	$e^{(0,9177 - 4,279 * A^2 - 1,38 * A^2^2)}$ (22)
$0,6 \leq A^2$	$e^{(1,2937 - 5,709 * A^2 + 0,0186 * A^2^2)}$ (23)

Etape 5: Interprétation

Lecture de la comparaison entre la valeur de la statistique calculée et la valeur critique

Résultat	Conclusion statistique	Décision
Valeur de la statistique > Valeur critique	On rejette H0	Les données ne suivent pas une loi normale

Lecture de la p-Value

Résultat	Conclusion statistique	Conclusion pratique
$P < \alpha\%$	On rejette H0	Les données ne suivent pas une loi normale au risque de se tromper de X%
$P > \alpha\%$	On retient H0	Les données suivent une loi normale au risque de se tromper de X%

Exemple : Pluie annuelle à la station de Batna (N=69, de 1936 à 2004).

```
library(goftest)
```

```
ad.test(HYPOBATNAan$Pan, "pnorm", mean=365.0663043, sd=105.3786855)
```

Anderson-Darling test of goodness-of-fit

Null hypothesis: Normal distribution

with parameters mean = 365.0663043, sd = 105.3786855

```
data: HYPOBATNAan$Pan
```

```
An = 0,44335, p-value = 0,8044 > 0,05, alors la série suit une loi normale.
```

Hydrolab 2010 donne

```
% U Anderson
```

```
= 0,343
```

La cellule est verte alors la série suit une loi normale.

Références bibliographiques

Anderson TW., Darling DA. (1954). A test of goodness of fit. *J. Am. Stat. Assoc.* 49, 765–769.

Wilcoxon F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin.* 1(6). p. 80-83. DOI: [10.2307/3001968](https://doi.org/10.2307/3001968), JSTOR 3001968

Mann HB. (1945). Nonparametric test against trend. *Econometrica* 13:245-259.

https://www.jstor.org/stable/pdf/1907187.pdf?seq=1#page_scan_tab_contents

Mann HB, Whitney DR. (1947). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.* 18(1). p. 50-60.

DOI: [10.1214/aoms/1177730491](https://doi.org/10.1214/aoms/1177730491)

Kendall MG. (1975). Rank correlation methods, 4th ed. Charles Griffin, London. UK.

Laborde JP., Mouhous N. (1998). Logiciel développé au CNRS, Université de Nice - Sophia Antipolis. France.

Pettitt AN. (1979). A non-parametric approach to the change-point problem. *Applied Statistics*, 28, 126-135. DOI: [10.2307/234672](https://doi.org/10.2307/234672)

Stephens MA. (1974). EDF Statistics for Goodness of Fit and Some Comparisons, *Journal of the American Statistical Association*, 69, pp. 730-737.

Stephens MA (1986). Tests based on EDF statistics. In: D'Agostino, R.B. and Stephens, M.A., eds.: Goodness-of-Fit Techniques. Marcel Dekker, New York.

Wald A., Wolfowitz J. (1940). On a test whether two samples are from the same population. *Ann. Math. Statist.*, 11, 147-162.

https://www.jstor.org/stable/2235872?seq=1#page_scan_tab_contents

WMO. (1988). Analyzing long time series of hydrological data with respect to climate variability. Project description, *WCAP-3, WMOTFD-No. 224*, Geneva, Switzerland.

<https://water.usgs.gov/osw/wcp-water/WCAP-3.pdf>