

# Recherche documentaire de tests d'adéquation

TRAN QUOC NHAT HAN - AMINE MOUSTATIH

8 octobre 2018

## Sommaire

<b>1</b>	<b>Test <math>\chi^2</math></b>	<b>3</b>
<b>2</b>	<b>Test de Shapiro-Wilk</b>	<b>3</b>
2.1	Histoire en bref . . . . .	3
2.2	Processus . . . . .	3
2.3	Code en R . . . . .	4
<b>3</b>	<b>Anderson Darling</b>	<b>5</b>

## Résumé

Ce rapport est un petit recherche documentaire concernant quelques tests d'adéquation usuels :

- Test  $\chi^2$
- Test de Shapiro-Wilk
- Test d'Anderson Darling

# 1 Test $\chi^2$

## 2 Test de Shapiro-Wilk

### 2.1 Histoire en bref

Ce test non-paramétrique est publié en 1965 par *Samuel Sanford Shapiro* et *Martin Wilk* pour tester si un échantillon d'une variable **continue** (sous 2000 observations) suit **une loi normale**.

### 2.2 Processus

Soit une variable continue  $X$  dont  $n$  observations étaient réalisées  $x_1, x_2, \dots, x_n$ .

Soient 2 hypothèses :

- $H_0$  :  $X$  suit une loi normale  $N(\mu, \sigma^2)$ .
- $H_1$  :  $X$  ne suit pas la loi normale.

Nous effectuons le test comme suivant :

1. Ordonner les réalisations dans l'ordre croissant  $y_1 \leq y_2 \leq \dots \leq y_n$ .
2. Calculer

$$S^2 = \sum_1^n (y_i - \bar{y})^2 = \sum_1^n (x_i - \bar{x})^2$$

Donc  $\bar{y}, \bar{x}$  désignent la moyenne de  $y, x$  respectivement.

3. Calculer des différences (entre le premier  $y_1$  et le dernier  $y_n$ , le deuxième  $y_2$  et l'avant-dernier  $y_{n-1}$ , et ainsi la suite, le médian  $y_{k+1}$  est ignoré si  $n = 2k + 1$ ). Appliquer un coefficient de pondération lu dans la table 1. Les additionner et élever au carré.

$$b^2 = \left( \sum_1^{\lfloor \frac{n}{2} \rfloor} a_i (y_{n+1-i} - y_i) \right)^2$$

4. Calculer le statistique du test

$$W = \frac{b^2}{S^2}$$

5. Rechercher la valeur  $W$  dans la table 2. Petit  $W$  signifie une distribution non normale. Choisir la valeur plus proche à  $W$  dans la ligne correspondant à  $n$ . Regarder alors le niveau (*level*) de signification  $p$ -value. Si  $p$ -value  $> \alpha$  ( $\alpha$  est le risque, souvent 1% ou 5%), l'hypothèse  $H_0$  est accepté.

$i \backslash n$	2	3	4	5	6	7	8	9	10
1	0.7071	0.7071	0.6872	0.6646	0.6431	0.6233	0.6052	0.5888	0.5739
2	—	.0000	.1677	.2413	.2806	.3031	.3164	.3244	.3291
3	—	—	—	.0000	.0875	.1401	.1743	.1976	.2141
4	—	—	—	—	—	.0000	.0561	.0947	.1224
5	—	—	—	—	—	—	—	.0000	.0399

FIGURE 1 – Coefficients  $a_{n+1-i}$  pour  $n = 1..10$

$n$	Level								
	0.01	0.02	0.05	0.10	0.50	0.90	0.95	0.98	0.99
3	0.753	0.756	0.767	0.789	0.959	0.998	0.999	1.000	1.000
4	.687	.707	.748	.792	.935	.987	.992	.996	.997
5	.686	.715	.762	.806	.927	.979	.986	.991	.993
6	0.713	0.743	0.788	0.826	0.927	0.974	0.981	0.986	0.989
7	.730	.760	.803	.838	.928	.972	.979	.985	.988
8	.749	.778	.818	.851	.932	.972	.978	.984	.987
9	.764	.791	.829	.859	.935	.972	.978	.984	.986
10	.781	.806	.842	.869	.938	.972	.978	.983	.986

FIGURE 2 – Pourcentage de  $W$  pour  $n = 1..10$

## 2.3 Code en R

```
shapiro.test(x)
```

$x$  désigne un vecteur de données.

### Exemples

```
> shapiro.test(rnorm(100, mean = 5, sd = 3))
```

Résultat de la commande:

Shapiro-Wilk normality test

```
data: rnorm(100, mean = 5, sd = 3)
```

```
W = 0.9895, p-value = 0.6211
```

$p$  est significativement plus grand que  $\alpha = 5\%$ . L'hypothèse nulle est donc non rejetable.

```
> shapiro.test(runif(100, min = 2, max = 4))
```

```
Résultat de la commande:  
Shapiro-Wilk normality test
```

```
data: runif(100, min = 2, max = 4)  
W = 0.9337, p-value = 8.077e-05
```

$p$  est trop faible. L'échantillon, par conséquence, ne suit pas la loi normale.

### 3 Anderson Darling

## Références

- [1] S. S. Shapiro et M. B. Wilk, An analysis of variance test for normality (complete samples)  
[https://github.com/haghighi/ST516/blob/master/Articles/Shapiro-Wilks%20test/An%20analysis%20of%20variance%20test%20for%20normality%20\(complete%20samples\).pdf](https://github.com/haghighi/ST516/blob/master/Articles/Shapiro-Wilks%20test/An%20analysis%20of%20variance%20test%20for%20normality%20(complete%20samples).pdf)
- [2] <http://www.jybaudot.fr/Inferentielle/testsnormalite.html>
- [3] <http://www.sthda.com/french/wiki/test-de-normalite-avec-r-test-de-shapiro-wilk>
- [4] <https://stat.ethz.ch/R-manual/R-patched/library/stats/html/shapiro.test.html>