

C. CARCASSONNE

**Angles de droits et de revers. Distribution circulaire**

*Mathématiques et sciences humaines*, tome 48 (1974), p. 29-41

[http://www.numdam.org/item?id=MSH\\_1974\\_\\_48\\_\\_29\\_0](http://www.numdam.org/item?id=MSH_1974__48__29_0)

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1974, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques  
<http://www.numdam.org/>

## ANGLES DE DROITS ET DE REVERS DISTRIBUTION CIRCULAIRE

par

C. CARCASSONNE <sup>1</sup>

### RÉSUMÉ

*Dans cet article, on traite un échantillon d'angles de droits et de revers de pièces de monnaies. On a cherché à en donner une description statistique correcte et à ajuster une loi théorique puis à construire un test d'homogénéité non paramétrique de deux échantillons distribués sur le cercle.*

### SUMMARY

*In this article, the author treats a sample of obverse and reverse of coins. We first try to give a correct statistical description of this sample, then we construct a distribution-free two-samples test of homogeneity on circle.*

L'échantillon dont nous présentons l'étude se trouve décrit dans un article de trois pages de E. Millau paru voici dix ans dans la *Revue belge de numismatique* [1].

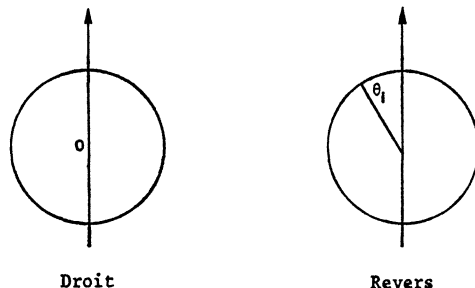
Cet échantillon est constitué de 225 pièces de monnaies qui appartiennent au Cabinet des Médailles de Bruxelles et qui frappées à Antioche portent lisiblement la marque de cet atelier. La technique de fabrication des monnaies anciennes étant mal connue, l'étude statistique de l'orientation des axes des deux côtés d'une pièce pourra peut-être donner une indication intéressante sur ce qui a été une des seules fabrications industrielles de l'Antiquité.

Décrivons, en deux mots, la frappe d'une pièce.

Un graveur incise les images choisies inversées, dans deux blocs de fer durci, ce sont les « coins », puis ces matrices sont enchassées dans deux plaques de métal. Un morceau de métal ou flan est placé entre le coin supérieur et le coin inférieur ou « pile ». Le monnayeur d'un vigoureux coup de marteau frappe le coin supérieur, le métal du flan reflue alors dans le creux des matrices et restitue ainsi le motif en relief. Le côté pile de la pièce s'appelle le *droit* et le côté supérieur le *revers* ou face.

La position relative des axes de droit et de revers varie d'une pièce à l'autre, ces variations correspondent à des différences dans la fabrication des coins ou dans la frappe elle-même.

L'axe du droit étant pris comme origine, l'angle formé avec lui par l'axe du revers a été mesuré. Les valeurs observées sont des points sur la circonférence d'un cercle. Notons  $\theta_i$ , l'angle formé entre l'axe origine et le vecteur joignant la  $i^{\text{ème}}$  observation au centre du cercle.



1. Centre de mathématique sociale EPHE, 6<sup>e</sup> section.

Les données sont résumées dans le tableau ci-dessous :

Intervalle de classe (en degrés)		Frequences observées
de	a	
307,5	312,5	1
312,5	317,5	1
317,5	322,5	3
322,5	327,5	9
327,5	332,5	9
332,5	337,5	9
337,5	342,5	10
342,5	347,5	16
347,5	352,5	9
352,5	357,5	11
357,5	2,5	13
2,5	7,5	5
7,5	12,5	4
12,5	17,5	2
17,5	22,5	2
22,5	27,5	0
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮
117,5	122,5	0
122,5	127,5	1
127,5	132,5	3
132,5	137,5	1
137,5	142,5	5
142,5	147,5	6
147,5	152,5	10
152,5	157,5	12
157,5	162,5	18
162,5	167,5	16
167,5	172,5	18
172,5	177,5	12
177,5	182,5	12
182,5	187,5	4
187,5	192,5	2
192,5	197,5	0
197,5	202,5	1
Total		225

Tableau I

L'histogramme des fréquences est construit en plaçant sur un cercle de rayon unité et à chaque milieu de classe, un bâtonnet proportionnel à l'effectif de la classe.

La répartition est très caractéristique, sa seule vue suffit à rejeter l'hypothèse d'une répartition uniforme.

Les problèmes que nous avons cherché à résoudre sont :

- I. Donner une description statistique correcte de cet échantillon.
- II. Ajuster une loi théorique.

Cette étude a été l'occasion d'une revue des méthodes utilisées pour aborder des problèmes comparables sur les variables circulaires. Nous en donnons un bref compte rendu.

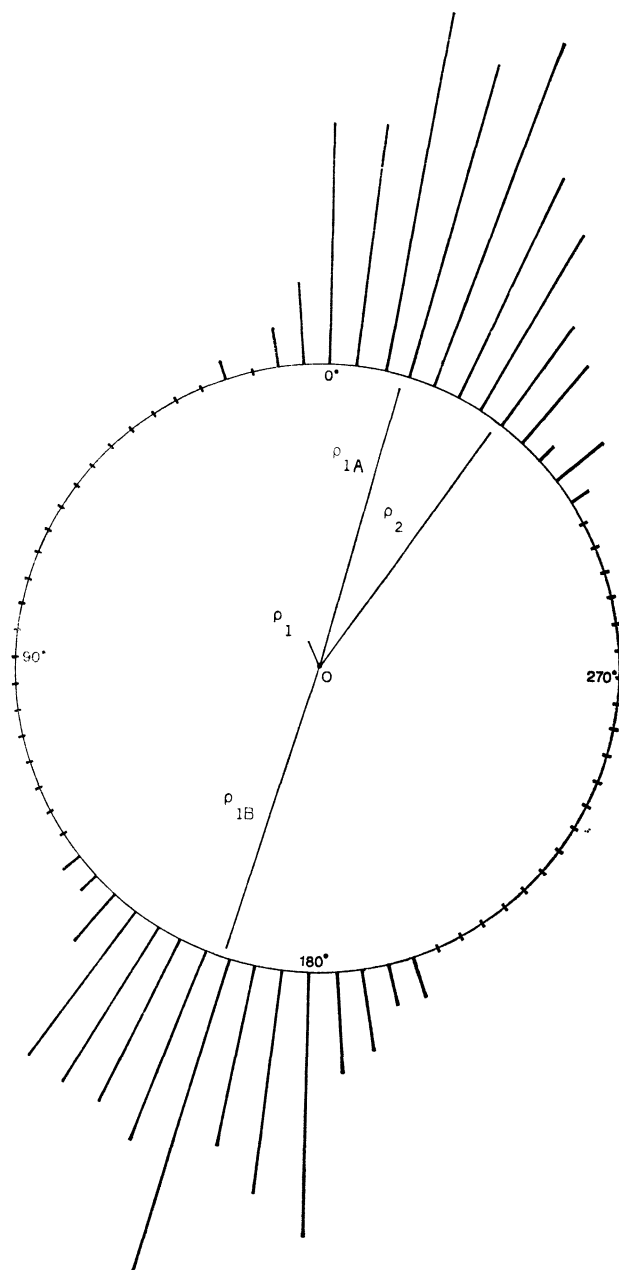


Fig. 2.

## 1. DESCRIPTION STATISTIQUE

Les résultats du calcul de la moyenne arithmétique et plus généralement de celui de tous les moments d'une variable circulaire dépendent de la manière dont ont été mesurés les angles et de l'origine choisie. Par exemple, la moyenne arithmétique des angles  $\theta_1 = 10^\circ$  et  $\theta_2 = 14^\circ$  est  $12^\circ$ ; si on décale l'origine de  $12^\circ$ , les deux angles  $\theta'_1 = 2^\circ$  et  $\theta'_2 = 358^\circ$  ont pour moyenne arithmétique  $180^\circ$ .

On ne peut donc utiliser la moyenne et la variance comme indicateurs de la valeur centrale et de la dispersion. Par contre, on peut utiliser dans ce but, les premiers coefficients de Fourier de la distribution empirique qui ne présentent pas ces inconvénients.

Soit :

$$c_1 = \frac{1}{\sum n_i} \sum_{i=1}^n n_i \cos \theta_i$$

$$s_1 = \frac{1}{\sum n_i} \sum_{i=1}^n n_i \sin \theta_i$$

Le vecteur  $(c_1, s_1)$  donne la position du centre de gravité des points observés. Si on passe en coordonnées polaires, on a :

$$c_1 = \rho_1 \cos \theta_0$$

$$s_1 = \rho_1 \sin \theta_0.$$

D'où  $\theta_0 = \text{Arctg} \frac{s_1}{c_1}$ , le signe de  $\theta_0$  étant déterminé d'après les signes de  $s_1$  et  $c_1$ .

$$\rho_1 = (c_1^2 + s_1^2)^{1/2}$$

L'angle  $\theta_0$  donne la direction moyenne des axes.

$\theta_0$  est un indicateur de valeur centrale, il joue un rôle analogue pour les distributions circulaires à celui que joue la moyenne d'une distribution linéaire.

Cette direction moyenne  $\theta_0$  est telle que la résultante des projections de tous les points observés sur cette direction est égale à  $\rho_1$  et sur la direction perpendiculaire, égale à 0. La valeur de  $\rho_1$  donne une indication sur la dispersion. Plus les observations sont groupées et plus grand est  $\rho_1$  :  $\rho_1$  est inversement proportionnel à la dispersion.  $\rho$  et  $\theta_0$  sont indépendants de l'origine choisie.

On peut prendre comme origine l'angle  $\theta_0 = \text{Arctg} \frac{s_1}{c_1}$ , les coefficients de Fourier valent alors respectivement :

$$c'_1 = \rho_1 \quad \text{et} \quad s'_1 = 0$$

On remarque que si la distribution est symétrique par rapport à cette position moyenne  $\theta_0$ , alors tous les coefficients  $s'_k$  :

$$s'_k = \frac{\sum_{i=1}^n n_i \sin k (\theta_i - \theta_0)}{\sum_{i=1}^n n_i}$$

sont nuls. S'il n'en est pas ainsi, on peut calculer  $(c'_2, s'_2)$ ;  $c'_2 = \rho_2 \cos \varphi$  et  $s'_2 = \rho_2 \sin \varphi$ , l'angle  $\varphi$  caractérise l'asymétrie de la distribution,  $\rho_2$  est inversement proportionnel à la dispersion, comme  $\rho_1$ .

D'autre part, on peut remarquer que si la distribution présente une symétrie par rapport au centre du cercle alors tous les  $(c_k, s_k)$  de rang impair sont nuls.

Dans notre cas l'histogramme suggère la présence d'une symétrie par rapport au centre et plus précisément deux axes de symétries perpendiculaires.

Dans le cas de notre échantillon, on a :

$$c_1 = 0,067$$

$$s_1 = 0,042$$

ce qui paraît vérifier l'hypothèse faite ci-dessus.

D'autre part :

$$c_2 = 0,65$$

$$\text{d'où } \varphi = 325^\circ$$

$$s_2 = -0,45$$

$$\rho_2 = 0,79$$

La vue de l'histogramme suggère de partager l'échantillon en deux sous-échantillons :

1. Sous-échantillon A : les revers sont plutôt dans le même sens que les droits;
2. Sous-échantillon B : les revers et les droits sont plutôt tête-bêche.

On obtient respectivement :

<i>Sous-échantillon A</i>	<i>Sous-échantillon B</i>
104 exemplaires.	121 exemplaires
$c_{1A} = 0,94$	$c_{1B} = -0,93$
$s_{1A} = -0,24$	$s_{1B} = 0,28$
$\theta_{0A} = \text{Arctg} \frac{-0,24}{0,94} = 360^\circ - 14^\circ = 346^\circ$	$\theta_{0B} = \text{Arctg} \frac{-0,93}{0,28} = 163^\circ$
$\rho_{1A} = 0,97$	$\rho_{1B} = 0,97$

En calculant tous les coefficients  $(c_k, s_k)$ , on obtient une série finie de nombres qui permettent de reconstituer la distribution observée.

## 2. QUEL MODÈLE CONSTRUIRE POUR CETTE DISTRIBUTION CARACTÉRISÉE PAR DEUX AXES DE SYMÉTRIE PERPENDICULAIRES?

Une distribution bimodale  $f(t)$  de période  $2\pi$  dont les deux modes sont diamétralement opposés. S'il en est ainsi  $c_1$  et  $s_1$  sont nuls.

Appliquons la méthode du maximum de vraisemblance.

Soit une distribution  $f(t_i)$  où  $t_i$  peut s'écrire  $t_i = \theta_i - \varphi$ .

On veut que le maximum de la fonction de vraisemblance  $L$  se trouve en  $\varphi$ , pour tout  $\varphi$  donné par l'équation :

$$\sum_{i=1}^n \sin 2(\theta_i - \varphi) = 0$$

$$L = \prod_{i=1}^n f(t_i) \quad \text{dans le cas d'observations indépendantes.}$$

$L$  est maximum lorsque sa dérivée s'annule. En dérivant on a :

$$\sum_{i=1}^n \frac{f'(\theta_i - \varphi)}{f(\theta_i - \varphi)} = 0$$

D'autre part :

$$\sum_{i=1}^n \sin 2(\theta_i - \varphi) = 0$$

Les deux sommes étant égales pour des valeurs quelconques de  $\theta_i$ , l'égalité doit donc se faire terme à terme.

D'où :

$$\sin 2(\theta_i - \varphi) = \frac{f'(\theta_i - \varphi)}{f(\theta_i - \varphi)}$$

où :

$$\sin 2t = \frac{f'(t)}{f(t)}$$

d'où :

$$d\left(\frac{1}{2} \cos 2t\right) = d(\text{Log } f(t))$$

$$a + \frac{1}{2} \cos 2t = b + \text{Log } f(t)$$

d'où :

$$f(t) = Ce^{k \cos 2t}$$

$C$  et  $k$  sont deux constantes liées par la condition :

$$\int_0^{2\pi} Ce^{k \cos 2t} dt = 1$$

d'où :

$$C = \frac{1}{\int_0^{2\pi} e^{k \cos 2t} dt}$$
$$= \frac{1}{2\pi I_0(k)} \quad \text{avec} \quad I_0(k) = \sum_{j=0}^{\infty} \frac{1}{(j!)^2} \left(\frac{k}{2}\right)^{2j}$$

$I_0(k)$  est la fonction de Bessel d'ordre 0 d'argument imaginaire pur. On a donc :

$$f(\theta) = \frac{e^{k \cos 2(\theta - \varphi)}}{2\pi I_0(k)}$$

Les points d'inflexion sont les solutions de :

$$f''(t) = 0$$

soit, en effectuant le calcul, les solutions de :

$$\cos 2t - k \sin^2 2t = 0;$$

cette équation a pour solutions :

$$\cos 2t = -\frac{1 \pm (1 + 4k^2)^{1/2}}{2k}.$$

seule  $\cos 2t = -\frac{1}{2k} + \left(1 + \frac{1}{4k^2}\right)^{1/2}$  est possible ; on en tire :  $2t = \text{Arccos } B$ .

D'où les 4 solutions :

$$t_1 = \frac{1}{2} \text{Arccos } B$$

$$t_2 = \pi - t_1$$

$$t_3 = -t_1$$

$$t_4 = (\pi + t_1)$$

Pour les grandes valeurs de  $k$  les points d'inflexion se trouvent approximativement à :

$$t_1 = \frac{1}{2\sqrt{k}}, \quad t_2 = \pi - \frac{1}{2\sqrt{k}}, \quad t_3 = -\frac{1}{2\sqrt{k}}, \quad t_4 = \pi + \frac{1}{2\sqrt{k}}$$

Si  $k$  est grand, son inverse influence la distribution comme un paramètre de dispersion.

Estimateurs de  $k$  et  $\varphi$ .

On a un échantillon  $\theta_1, \dots, \theta_n$  et on cherche à estimer les valeurs de  $k$  et  $\varphi$  par la méthode du maximum de vraisemblance et

$$L = \frac{1}{(2\pi)^n} \frac{1}{(I_0(k))^n} e^{-k \sum_{i=1}^n \cos 2(\theta_i - \varphi)}$$

$$\frac{\partial \text{Log } L}{\partial k} = -n \frac{I_0'(k)}{I_0(k)} + \sum_{i=1}^n \cos(\theta_i - \varphi)$$

$$\frac{\partial \text{Log } L}{\partial \varphi} = k \sum_{i=1}^n 2 \sin(\theta_i - \varphi)$$

$L$  est maximum pour

$$\sum_{i=1}^n \sin 2(\theta_i - \varphi) = 0 \quad \text{soit : } s'_2 = 0$$

et :

$$\frac{I_0'(k)}{I_0(k)} = \frac{1}{n} \sum_{i=1}^n \cos 2(\theta_i - \varphi) = c'_2$$

comme  $s'_2 = 0$ ,  $c'_2$  est égal au vecteur moyen.

On a donc :

$$\frac{I_0'(k)}{I_0(k)} = \rho_2.$$

On peut donc estimer  $\varphi$  et  $k$  à partir des observations. Le calcul de  $k$  se fait par interpolation à l'aide d'une table de Bessel [2] en remarquant que  $I_0'(k) = I_1(k)$ .

Par exemple ici on a :

$$c_2 = 0,65 \quad s_2 = -0,42 \quad \rho_2 = 0,79$$

donc :

$$\frac{I_1(k)}{I_0(k)} = 0,79$$

On trouve dans la table :

$k$	$e^{-k} I_0(k)$	$e^{-k} I_1(k)$	$\frac{I_1(k)}{I_0(k)}$
2,7	0,25818	0,20269	0,7851
2,8	0,25280	0,20073	0,7940

par interpolation linéaire on tire :

$$k = 2,7 + \frac{0,79 - 0,7851}{0,7940 - 0,7851} \times 0,1$$

$$k = 2,7 + 0,0449 = 2,7449$$

$$k \simeq 2,75$$



Cas limite :

$$f(t) = \frac{e^{k \cos 2(\theta - \varphi)}}{2\pi I_0(k)}$$

Lorsque  $k$  est grand les points d'inflexion se trouvent en :  $\frac{1}{2\sqrt{k}}$  ,  $\pi - \frac{1}{2\sqrt{k}}$  ,  $-\frac{1}{2\sqrt{k}}$  et  $\pi + \frac{1}{2\sqrt{k}}$  ;

( $k$  influence la distribution comme un facteur de dispersion). Supposons qu'il en soit ainsi, la distribution est très concentrée autour de  $\pi$  et  $\varphi + \pi$ . Introduisons un aléa numérique réduit en écrivant :

$$\theta - \varphi = \frac{t}{2\sqrt{k}}$$

la distribution de  $t$  est :

$$\begin{aligned} f(\tau) &= \text{Cte} \times e^{k \cos(t/2\sqrt{k})} \\ &= \text{Cte} \times e^{k \cos \frac{t}{\sqrt{k}}} \quad \text{ou} \quad e^{k \cos \frac{t'}{\sqrt{k}}} \quad \text{pour } t' = t + \pi; \end{aligned}$$

puisque l'argument du *cosinus* est petit, nous pouvons négliger les grandes puissances de  $t$  et écrire :

$$\cos \frac{t}{\sqrt{k}} \sim 1 - \frac{t^2}{k} + \dots$$

d'où :

$$f(t) = \text{Cte} e^{-\frac{kt^2}{k}} = \text{Cte} e^{-t^2}$$

$$f(t') = \text{Cte} e^{-\frac{t'^2}{2}}$$

La distribution se partage en deux morceaux l'un centré autour de 0, l'autre autour de  $\pi$ , chacun tendant vers une loi normale linéaire, d'écart-type  $\sigma = \frac{1}{2}$  et de moyenne 0 ou  $\pi$ .

Si  $k$  tend vers l'infini alors la distribution est entièrement groupée aux points 0 et  $\pi$  du cercle unité.

Dans tout ce qui précède, nous avons utilisé la méthode utilisée par R. von Mises pour construire une loi circulaire unimodale qu'il appelle lui-même loi normale circulaire dont la densité est la suivante :

$$f(\alpha) = \frac{e^{\alpha \cos(\alpha - \alpha_0)}}{2\pi I_0(k)} \quad 0 < \alpha < \pi$$

où  $I_0(k)$  est la fonction de Bessel de première espèce d'argument imaginaire pur. Cette loi est la plupart du temps citée sous le nom de loi de von Mises [3] [4].

Cette fonction a un mode pour  $\alpha = \alpha_0$  et un minimum pour  $\alpha_0 + \pi$ .

Le quotient entre les densités au mode et au minimum est :

$$\frac{f(\alpha_0)}{f(\alpha_0 + \pi)} = e^{2k};$$

pour  $k = 0$  la distribution dégénère et on a la distribution uniforme circulaire.

Pour  $k$  grand son inverse  $\frac{1}{k}$  influence la distribution normale circulaire de la même manière que  $\sigma^2$  influence la loi normale linéaire. Cette analogie conduit au second cas limite.

$$\text{Posons : } x = \sqrt{k}(\alpha - \alpha_0) \quad \frac{x}{\sqrt{k}} = \alpha - \alpha_0$$

$$f(x) = \text{Cte} \times e^{k \cos\left(\frac{x}{\sqrt{k}}\right)}$$

En développant on a :  $\cos\left(\frac{x}{\sqrt{k}}\right) = \frac{-x^2}{2k} + 0\left(\frac{1}{k}\right)$

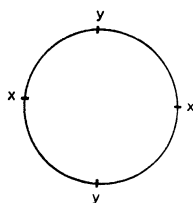
$$f(x) = \text{Cte} \times e^{-k\frac{x^2}{2k}} = \text{Cte} \times e^{-\frac{x^2}{2}}$$

On trouve la loi normale centrée réduite.

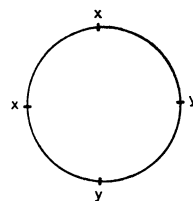
*Test* : Nous avons essayé, à la suggestion de M. Guilbaud, de transposer le test de Wilcoxon sur le cercle, pour tester l'homogénéité de deux échantillons observés. Soient deux échantillons : les  $x$  qui comptent  $n$  observations  $x_1, x_2, \dots, x_n$  et les  $y$  qui comptent  $m$  observations  $y_1, y_2, \dots, y_m$ ; les  $x$  et les  $y$  sont des variables circulaires. On peut construire un test non paramétrique d'homogénéité sur le cercle analogue au test de Wilcoxon sur un segment.

Le test ne porte pas sur les vraies valeurs des variables mais seulement sur les positions relatives des couples de  $x$  et des couples de  $y$ . Ce test repose sur l'idée que si deux échantillons proviennent de la même population alors les observations en  $x$  et en  $y$  doivent être bien enchevêtrées; on dit qu'il y a :

enchevêtrement si on a :



qu'il n'y en a pas si on a :



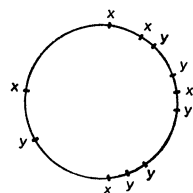
Prenons un exemple fictif de 5  $x$  et 6  $y$ .

$$x_1 = 96^\circ, \quad x_2 = 70^\circ, \quad x_3 = 180^\circ, \quad x_4 = 275^\circ, \quad x_5 = 17^\circ \quad \text{et}$$

$$y_1 = 0^\circ, \quad y_2 = 27^\circ, \quad y_3 = 60^\circ, \quad y_4 = 210^\circ, \quad y_5 = 280^\circ, \quad y_6 = 300^\circ$$

Portons les points sur le cercle en oubliant les indices;

nous avons le rangement :



Quels sont les rangements possibles? Comment se place l'observation dans ces possibles?

Dressons la liste des possibles, il y en a 42. On peut calculer facilement le nombre total de rangements quand  $n$  et  $m$  sont premiers entre eux. Il suffit alors de diviser par  $n + m$  le nombre de façons de choisir  $n$  éléments parmi  $n + m$ . Quand ce n'est pas le cas c'est plus difficile, car il y a des périodicités dans certains arrangements, mais cependant réalisable pour les faibles valeurs de  $m$  et  $n$ .

Liste des possibles (coupés arbitrairement en un point du cercle) :

	<i>I</i>		<i>I</i>
<i>x x x x x . . . . .</i>	0	<i>x x . x x . x . . . .</i>	46
<i>x x x x . x . . . . .</i>	20	<i>x x . x x . . x . . .</i>	54
<i>x x x x . . x . . . .</i>	32	<i>x x . x x . . . x . .</i>	54
<i>x x x x . . . x . . .</i>	36	<i>x x . x x . . . . x .</i>	46
<i>x x x x . . . . x . .</i>	32	<i>x x . x . x x . . . .</i>	52
<i>x x x x . . . . . x .</i>	20	<i>x x . x . x . x . . .</i>	62
<i>x x x . x x . . . . .</i>	30	<i>x x . x . x . . x . .</i>	64
<i>x x x . x . x . . . .</i>	44	<i>x x . x . x . . . x .</i>	58
<i>x x x . x . . x . . .</i>	50	<i>x x . x . . x x . . .</i>	62
<i>x x x . x . . . x . .</i>	48	<i>x x . x . . . x . . .</i>	66
<i>x x x . x . . . . x .</i>	38	<i>x x . x . . . . x . .</i>	62
<i>x x x . . x x . . . .</i>	48	<i>x x . x . . . x x . .</i>	60
<i>x x x . . x . x . . .</i>	56	<i>x x . x . . . x . x .</i>	58
<i>x x x . . x . . x . .</i>	56	<i>x x . . x x . . x . .</i>	64
<i>x x x . . x . . . x .</i>	48	<i>x x . . x x . . . x .</i>	60
<i>x x x . . . x x . . .</i>	54	<i>x x . . . x x . . . .</i>	62
<i>x x x . . . x . x . .</i>	56	<i>x x . . . x . x . . .</i>	68
<i>x x x . . . x . . x .</i>	50	<i>x x . . . x . x . . x .</i>	66
<i>x x x . . . . x x . .</i>	48	<i>x x . . . . x . x . .</i>	64
<i>x x x . . . . . x . x</i>	44	<i>x x . . . . x . x . x .</i>	62
<i>x x x . . . . . x x .</i>	30	<i>x . x . x . x . x . .</i>	70

2 100

Repérons l'enchevêtrement des *x* et des *y* par le nombre total d'enchevêtrements *I* entre les couples de *x* et de *y*.

Il y a ici :  $\frac{n(n-1)}{2} = \frac{5 \times 4}{2} = 10$  couples de *x*

et :  $\frac{m(m-1)}{2} = \frac{6 \times 5}{2} = 15$  couples de *y*.

donc :  $10 \times 15 = 150$  comparaisons à faire.

Dressons la courbe cumulée *F(I)* des rangements possibles par enchevêtrement croissant.

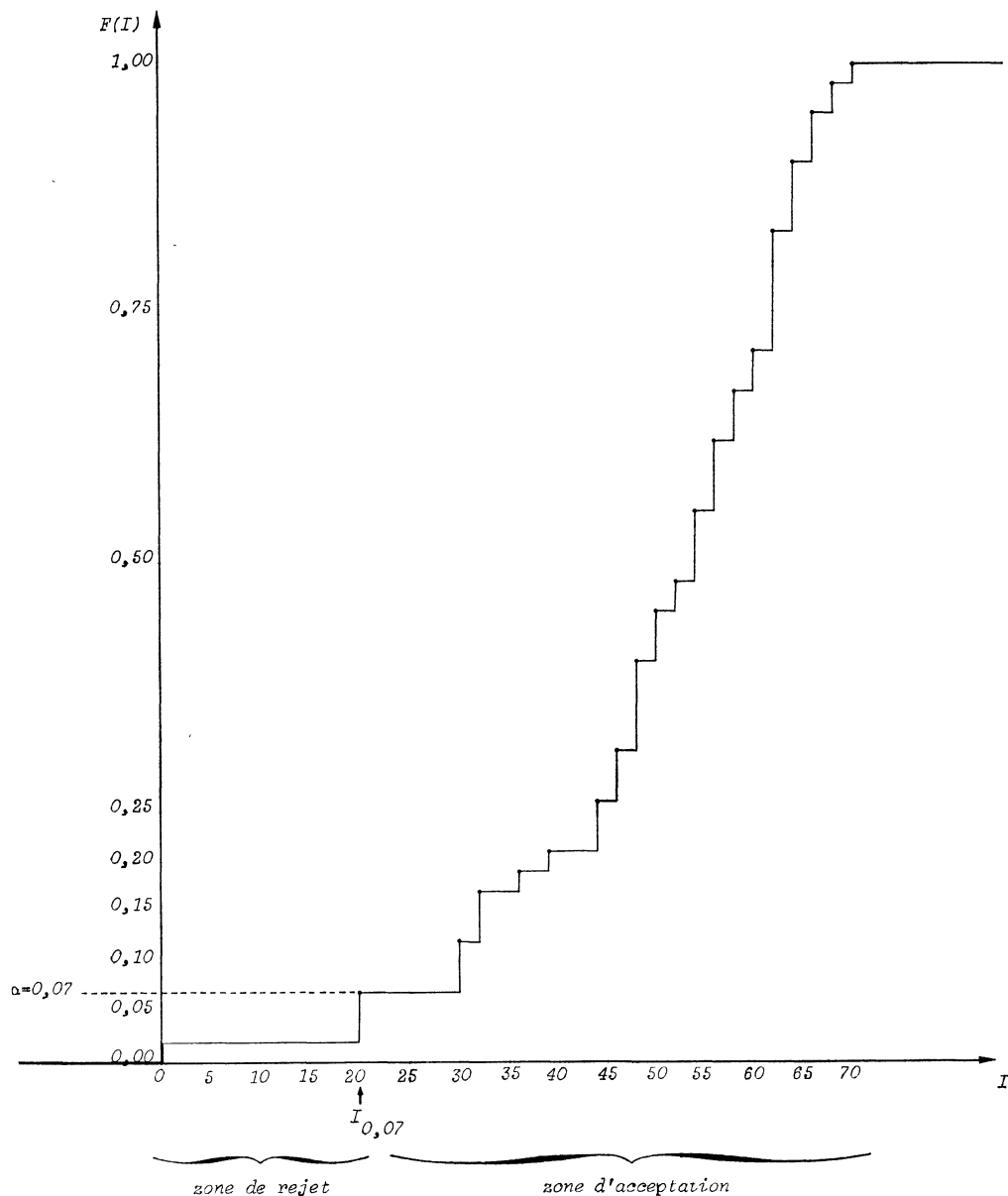
Tableau 2. *F(I)* :

<i>I</i>	<i>n<sub>I</sub></i>	$\Sigma n_I$	<i>F(I)</i>
0	1	1	0,02
20	2	3	0,07
30	2	5	0,12
32	2	7	0,17
36	1	8	0,19
38	1	9	0,21
44	2	11	0,26
46	2	13	0,31
48	4	17	0,40
50	2	19	0,45
52	1	20	0,48
54	3	23	0,55
56	3	26	0,62
58	2	28	0,67
60	2	30	0,71
62	5	35	0,83
64	3	38	0,90
66	2	40	0,95
68	1	41	0,98
70	1	42	1,00

Appelons *H*<sub>0</sub> l'hypothèse d'homogénéité des deux échantillons. Si l'observation correspond à un *F(I)* faible on considère que c'est un résultat rare quand *H*<sub>0</sub> est vraie et on rejettera *H*<sub>0</sub>. Le test est unilatéral, il prescrit de rejeter

$H_0$  dès que le nombre des enchevêtrements est dépassé, par une borne  $I_\alpha$ , la borne  $\alpha$  étant choisie de telle sorte que sous  $H_0$ ,  $F(I_\alpha) \leq \alpha$ .

Fixons par exemple  $\alpha = 0,10$ . Il y a 42 arrangements sur le cercle, donc  $F(I_\alpha) \leq \alpha$  pour les  $\frac{10}{100} \times 42$  arrangements de plus faible nombre d'enchevêtrements. Pour ces 4 premiers arrangements,  $I_\alpha = 30$ . Deux arrangements comportent 30 enchevêtrements, le seuil  $\alpha = 0,10$  est donc dépassé; pour  $I = 30$ ,  $F(I) = 0,12$ . On se ramène alors à  $I = 20$  pour lequel  $F(I) = 0,07$ . La probabilité du test n'est pas exactement 10 % mais 0,07 légèrement inférieure à 10 %.



### Moyenne et variance de $I$ .

Posons  $z_{ij, kl} = 1$  si les couples  $(x_i, x_j)$  et  $(y_k, y_l)$  sont enchevêtrés,  
 $= 0$  sinon.

Dans le cas où les points  $x_i, x_j, y_k, y_l$  sont indépendants et de même distribution, les valeurs 0 et 1 ont respectivement les probabilités  $\frac{2}{3}$  et  $\frac{1}{3}$ . On a  $E(z_{ij, kl}) = \frac{1}{3}$ .

Comme :

$$I = \sum_{i < j}^n \sum_{k < l}^m (z_{ij, kl})$$

$$E(I) = \frac{n(n-1)}{2} \times \frac{m(m-1)}{2} \times \frac{1}{3}$$

$$\text{Var } I = E[(I - E(I))]^2$$

$$\text{Var } I = E\left(\left[\sum (z_{ij, kl} - \frac{1}{3})\right]^2\right)$$

$$\text{Var } I = \sum E\left(z_{ij, kl} - \frac{1}{3}\right)\left(z_{gh, st} - \frac{1}{3}\right)$$

les termes tels que  $i \neq j \neq g \neq h$  et  $l \neq k \neq s \neq t$  sont nuls, car  $z_{ij, kl}$  et  $z_{gh, st}$  sont indépendants et ont pour moyenne  $\frac{1}{3}$ . Les termes dont les indices du terme de gauche sont égaux à ceux du deuxième terme sont au nombre de  $\frac{m(m-1)}{2} \times \frac{n(n-1)}{2}$ . Ils ont pour espérance la variance de  $z_{ij, kl}$  soit  $\frac{1}{3} \times \frac{2}{3} = \frac{2}{9}$ . Les termes tels que un des indices diffère sont au nombre de  $\frac{m(m-1)}{2} \times \frac{n(n-1)}{2} \times 2 \times (m+n-4)$ . Ils ont pour espérance Cov  $(z_{ij, kl}, z_{ij, sl})$  soit :

$$E(z_{ij, kl} \times z_{ij, sl}) - \frac{1}{3} \times \frac{1}{3} = 1 \times \frac{1}{6} - \frac{1}{9} = \frac{1}{18}.$$

Les termes tels que deux des indices diffèrent sont de deux espèces.

1<sup>er</sup> cas. — Les deux  $x$  ou les deux  $y$  sont différents. Il y a :

$$\frac{m(m-1)}{2} \times \frac{n(n-1)}{2} \times \left[ \frac{(m-2)(m-3)}{2} + \frac{(n-2)(n-3)}{2} \right] \text{ cas.}$$

On a l'espérance :  $\frac{8}{60} - \frac{1}{9} = \frac{4}{180}$ .

2<sup>e</sup> cas. — Un  $x$  et un  $y$  sont différents. il y a :  $\frac{m(m-1)}{2} \times \frac{n(n-1)}{2} \times 4 \times (m-2)(n-2)$  cas.

On a l'espérance :  $\frac{7}{60} - \frac{1}{9} = \frac{1}{180}$ .

Les termes tels que trois indices diffèrent ont une espérance nulle. On a donc en regroupant ces résultats.

$$\text{Var } I = \frac{n(n-1)}{2} \times \frac{m(m-1)}{2} \times \left[ \frac{2}{9} + 2 \times (m+n-4) \times \frac{1}{18} + \left[ \frac{(m-2)(m-3)}{2} + \frac{(n-2)(n-3)}{2} \right] \times \frac{4}{180} \right. \\ \left. + 4(m-2)(n-2) \times \frac{1}{180} \right] = \frac{nm(n-1)(m-1)(m+n)(m+n+1)}{360};$$

dans le cas :  $m = 6, n = 5$  on a :

$$E(I) = 50$$

$$\text{Var } I = 220$$

$$\sigma(I) = 14,8$$

Divers tests ont été construits pour les distributions circulaires. On peut citer en particulier Watson [5] qui donne un test d'homogénéité non paramétrique de deux échantillons. Le test est basé sur la valeur de la statistique :

$$R_N^{2*} = \sum_{j=1}^m \sum_{k=1}^n \cos 2\pi \left\{ (r_j - r_k)/N \right\}$$

où  $m$  et  $n$  sont les effectifs des échantillons et  $r_j$  et  $r_k$  les rangs des observations comptés à partir d'une origine arbitraire. De grandes valeurs de  $R^*_N$  suggèrent une différence entre les deux populations; le test peut être utilisé pour tester l'axe de symétrie, en relevant les points en dessous de l'axe sur la partie au-dessus.

On trouve dans Schach [6] et Beran [7] d'autres tests non paramétriques de l'axe de symétrie d'une distribution circulaire et de l'uniformité.

Dans Stephens [8] [9] on trouve des tests adaptés plus particulièrement à la distribution de von Mises.

## BIBLIOGRAPHIE

- [1] MILLAU E., « Angles de droits et de revers », *Revue belge de numismatique*, t. CVI, 1960.
- [2] ABRAMOVITZ M., ed., *Handbook of mathematical functions with formulars graphs and mathematical tables*, New York, Dover, 1965.
- [3] GUMBEL E. J., GREENWOOD J. A., DURAND D. « The circular normal distribution : Theory and tables », *J. Am. statist. Ass.*, 1954, t. 48, p. 131-152.
- [4] GUMBEL E. J., « Applications of the circular normal distribution », *J. Am. statist. Ass.*, 1954, t. 49.
- [5] WATSON G. S. et WHEELER S., « A distribution free two sample tests on circle », *Biometrika* 1964, 51, p. 256.
- [6] SCHACH S., « Non parametric symmetry tests for circular distribution », *Biometrika* 1969, 56-3, p. 571.
- [7] BERAN R. J., « The derivation of non parametric two sample tests from tests for uniformity of a circular distribution », *Biometrika* 56, 3, p. 561.
- [8] STEPHENS M. A., « A goodness of fit statistic for the circle with some comparison », *Biometrika* 1969, 56, 1, p. 161.
- [9] STEPHENS M. A., « Tests for the von Mises distribution », *Biometrika* 56, 1, p. 149.