

REVUE DE STATISTIQUE APPLIQUÉE

J. RAISON

Le test des signes. Note complémentaire théorique

Revue de statistique appliquée, tome 5, n° 3 (1957), p. 91-98

http://www.numdam.org/item?id=RSA_1957__5_3_91_0

© Société française de statistique, 1957, tous droits réservés.

L'accès aux archives de la revue « *Revue de statistique appliquée* » (<http://www.sfds.asso.fr/publicat/rsa.htm>) 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/>

LE TEST DES SIGNES

NOTE COMPLÉMENTAIRE THÉORIQUE

par
J. RAISON

Comme l'indique J. HEMELRIJK (1), le test des signes est sans doute le test dont l'existence est la plus ancienne puisque J. ARBUTHNOT l'appliquait sous une forme rudimentaire dès 1710 (2).

Le test des signes est utilisé pour tester l'hypothèse H_0 , que deux variables stochastiques indépendantes \underline{x} et \underline{y} ont la même distribution, lorsque l'on dispose de n paires d'observations indépendantes (x_i, y_i) de ces deux variables, observations faites dans les mêmes conditions pour les deux éléments de chaque paire et pouvant varier de paire à paire. La distribution des observations pouvant être d'une part différente de la distribution normale, on ne peut utiliser le test de Student ou d'autres tests dits paramétriques. Les x_i, y_i pouvant, d'autre part, être difficiles à déterminer, il n'est pas toujours possible de les quantifier.

Le test des signes est basé sur le nombre des signes (+ ou -) des différences $x_i - y_i$ des deux observations de chacune des paires (3). Il consiste à tester l'hypothèse que chacune de ces différences fait partie d'une distribution à médiane nulle. On rejette l'hypothèse dans le cas où les nombres de signes + et - diffèrent de façon significative.

W. J. DIXON et A. M. MOOD proposaient en 1946 de compter le nombre d'égalités pour moitié en différences positives et pour moitié en différences négatives. W. J. DIXON et F. J. MASSEY Jr précisent dans leur ouvrage "Introduction to statistical analysis" (1951) qu'il faut exclure les cas d'égalité. J. HEMELRIJK a démontré que le test était beaucoup plus puissant si l'on omettait les égalités au lieu de les diviser.

1. - a) En admettant que pour toute paire d'observations (x_i, y_i) il y ait une probabilité p pour que $x_i > y_i$, on aura pour les n différences $(x_i - y_i)$ une probabilité :

$C_n^k p^k (1-p)^{n-k}$ de trouver k différences positives et une probabilité

$C_n^k p^{n-k} (1-p)^k$ de trouver k différences négatives. (On a une distribution binominale p dans le 1er cas et $1-p$ dans le second), k différences négatives correspondant d'ailleurs à $n-k$ différences positives et $C_n^k = C_n^{n-k}$.

(1) J. HEMELRIJK - A theorem on the sign test when ties are present. Proceedings Nederl. Akademie van Wetenschappen Series A 55 (1952) p. 322.

(2) I. A. TODHUNTER - A history of the mathematical theory of probability - Cambridge and London 1865 p. 131, cité par J. NEYMAN, H. CRAMER et J. HEMELRIJK.

H. FREUDENTHAL - La naissance de la pensée statistique dans les sciences modernes. Conférences de Physique Nouvelle série 29, 57-65 (1951), cité par J. HEMELRIJK.

(3) W. J. DIXON et A. M. MOOD - The statistical sign test, Journal of the American statistical association 41, 557-566 (1946).

La probabilité d'obtenir k différences positives ou négatives est donc égale à

$$C_n^k [p^k (1-p)^{n-k} + p^{n-k} (1-p)^k]$$

pour toutes les valeurs de k inférieures à $\frac{n-1}{2}$ si k est impair et à $\frac{n}{2}$ si k est pair. Dans ce dernier cas, la probabilité d'obtenir $\frac{n}{2}$ différences positives et par suite $\frac{n}{2}$ différences négatives est égale à

$$C_n^{\frac{n}{2}} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}}$$

Si \underline{x} et \underline{y} sont indépendantes et ont la même distribution $\underline{z} = \underline{x} - \underline{y}$ a la même distribution de probabilité que $-\underline{z} = \underline{y} - \underline{x}$. Ceci revient à dire que la probabilité d'avoir $\underline{z} > 0$ est égale à la probabilité d'avoir $\underline{z} < 0$.

$$\begin{aligned} P[\underline{z} > 0] &= P[\underline{z} < 0]; \\ \text{Si } P[\underline{z} = 0] = 0 \text{ on a alors } P[\underline{z} > 0] &= P[\underline{z} < 0] = 1/2 \\ \text{Si } P[\underline{z} = 0] > 0 \text{ on a } P[\underline{z} > 0 | \underline{z} \neq 0] &= \frac{P[\underline{z} > 0 \wedge \underline{z} \neq 0]}{P[\underline{z} \neq 0]} \end{aligned}$$

par application de la formule de la probabilité conditionnelle de A sous la condition B

$$P[A | B] = \frac{P[A \wedge B]}{P[B]}$$

$$\frac{P[\underline{z} > 0 \wedge \underline{z} \neq 0]}{P[\underline{z} \neq 0]} = \frac{P[\underline{z} > 0]}{P[\underline{z} < 0 \vee \underline{z} > 0]} = \frac{P[\underline{z} > 0]}{P[\underline{z} > 0] + P[\underline{z} < 0]} = \frac{1}{2}$$

Pour $P[\underline{z} < 0 | \underline{z} \neq 0]$ il en est de même (x 4).

Si \underline{x} et \underline{y} sont indépendants et ont la même distribution, la probabilité p pour que $\underline{x} - \underline{y}$ soit > 0 sous la condition que $\underline{x} - \underline{y}$ soit $\neq 0$ est donc égale à $1/2$.

b) Inversement, pour tester l'hypothèse $p = 1/2$ que \underline{x} et \underline{y} ont la même distribution, on pourra prendre le nombre des différences, positives par exemple, soit k et accepter ou rejeter l'hypothèse suivant que k se rapproche ou s'éloigne de la valeur théorique $\frac{1}{2}n$. k a une répartition binominale. Il faut donc déterminer les zones critiques de sa distribution qui nous indiqueront à un niveau de confiance donné α , s'il faut accepter ou rejeter cette hypothèse $p = \frac{1}{2}$.

La probabilité pour que l'on ait $0, 1, 2, \dots, k$ différences positives si leur probabilité est p est donnée par la somme des coefficients du développement du binôme.

$$\sum_{i=0}^k C_n^i p^i (1-p)^{n-i}$$

qu'on peut transformer en

$$\sum_{i=0}^k C_n^i p^i (1-p)^{n-i} = 1 - \frac{\int_0^p x^k (1-x)^{n-k-1} dx}{\int_0^1 x^k (1-x)^{n-k-1} dx}$$

et, en posant $k = r - 1$:

$n - k - 1 = s - 1$, on obtient :

$$\sum_{i=0}^k C_n^i p^i (1-p)^{n-i} = 1 - \frac{\int_0^p x^{r-1} (1-x)^{s-1} dx}{\int_0^1 x^{r-1} (1-x)^{s-1} dx}$$

En notant $B_p(r, s) = \int_0^p x^{r-1} (1-x)^{s-1} dx$, pour représenter la fonction Bêta incomplète, on a :

$$\sum_{i=0}^k C_n^i p^i (1-p)^{n-i} = 1 - \frac{B_p(r, s)}{B(r, s)}$$

En utilisant la notation de K. PEARSON dans "Tables of the Incomplete Beta function" (London 1934) et en posant :

$$I_p(r, s) = \frac{B_p(r, s)}{B(r, s)}$$

On obtient :

$$\sum_{i=0}^k C_n^i p^i (1-p)^{n-i} = 1 - I_p(k+1, n-k) = I_{1-p}(n-k, k+1)$$

ou, en remplaçant k par $\ell - 1 =$

$$\sum_{i=\ell}^0 C_n^i p^i (1-p)^i = I_p(\ell, n-\ell+1) = 1 - I_{1-p}(n-\ell+1, \ell)$$

Les tables $I_p(r, s)$ tabulées par PEARSON le sont pour r et s de 1 à 50 ; on peut donc aller jusqu'à $n = 99$.

H.O. BARTLEY et E.R. FITCH ont réussi à faire figurer sur une page un nomogramme de $I_p(r, s)$ pour $r \leq 200$ et $s \leq 60$: "A chart for the incomplete Beta function and the cumulative binominal distribution" *Biometrika* 38, (1951) pp. 423-428.

À défaut de ces tables cumulatives, les tables du National Bureau of Standards 1950 "Tables of the binominal probability distribution" donnent la valeur des probabilités $C_n^k p^k (1-p)^{n-k}$ ainsi que les sommes cumulées pour $p = 0,01$ à $0,50$ et $n \leq 49$ avec 7 décimales. Les "Binominal tables" de ROMIG (1953) donnent ces valeurs pour $50 \leq n \leq 100$.

Lorsque $p = \frac{1}{2}$ les coefficients sont de la forme $2^{-n} C_n^k$.

La probabilité d'avoir 0, 1, 2...k différences positives est :

$$2^{-n} \sum_{i=0}^k C_n^i$$

A. Van WIJNGAARDEN a tabulé ces coefficients : *Proceedings Kon Nederl. Akademie Van Wetenschappen* 53 (1950) pp. 857-868 "Table of the cumulative symmetric binomial distribution".

Les zones critiques employées dans le cas du test des signes sont de la forme habituelle, soit bilatérale :

$$Z : \left| x - \frac{1}{2}n \right| \geq a$$

ce qui veut dire que l'on rejette l'hypothèse $H_0 : p = 1/2$ au cas où la différence entre le résultat trouvé x_0 (le nombre de succès ou différences positives dans n expériences et $1/2$ diffère de plus de a ,

soit unilatérales :

$$\text{droite } Z_d : x - \frac{1}{2}n \geq b$$

$$\text{gauche } Z_g : x - \frac{1}{2}n \leq -b$$

la probabilité d'apparition $1 - \beta(H_0)$ d'un résultat dans ces zones étant inférieure à une limite α seuil de confiance

$$1 - \beta(H) \leq \alpha$$

Dans le cas de la zone critique bilatérale, on aura alors :

$$1 - \beta(H_0) = 2^{-n} \sum_{i=0}^{\frac{1}{2}n-a} C_n^i + 2^{-n} \sum_{i=\frac{1}{2}n+a}^n C_n^i = 2^{-n+1} \sum_{i=0}^{\frac{1}{2}n-a} C_n^i$$

Pour les zones critiques unilatérales on aura de même

$$1 - \beta_d (H_0) = 1 - \beta_g (H_0) = 2^{-n} \sum_{i=0}^{\frac{1}{2}n-b} C_n^i$$

a et b sont choisis de façon à ce que $\frac{1}{2}n - a$ et $\frac{1}{2}n - b$ soient des nombres entiers
 $1 - \beta (H_0)$ est la limite de confiance réelle.

Il faut donc déterminer le plus grand a ou b répondant aux conditions données.
 Les tables indiquées peuvent être utilisées.

W. J. DIXON et A. M. MOOD ont calculé les valeurs $\frac{1}{2}n - a$ pour :

$$n \leq 100 \quad \text{pour } \alpha = 0,01 \quad 0,05 \quad 0,1 \quad \text{et} \quad 0,25$$

Voir DIXON et MASSEY "Introduction to statistical analysis (1951) table 10 p. 324.

Si l'on ne dispose pas de ces tables, on peut faire utilement emploi de l'approximation de la loi normale ($\mu = \frac{1}{2}n$; $\sigma^2 = \frac{1}{4}n$) pour trouver a et b. Il faut faire attention à ne pas oublier la correction de continuité $\frac{1}{2}$.

Variable déduite $y_{\frac{1}{2}\alpha}$ telle que $\int_{-\infty}^{y_{\frac{1}{2}\alpha}} e^{-\frac{u^2}{2}} du = \frac{1}{2}\alpha$ (dans le cas de la zone critique bilatérale) puis :

$$\frac{x_{\frac{1}{2}\alpha} - \frac{1}{2}n}{\frac{1}{2}\sqrt{n}} = y_{\frac{1}{2}\alpha} \quad \text{et enfin} \quad \frac{1}{2}n - a = x_{\frac{1}{2}\alpha} - \frac{1}{2}$$

J. HEMELRIJK a procédé à ce calcul pour $n \leq 100$ et $\alpha = 0,05$. Il a constaté pour toutes les valeurs de n, sauf $n = 17$ et $n = 94$, que les résultats obtenus par la méthode approchée étaient corrects. Pour les deux valeurs indiquées, la zone critique bilatérale trouvée était erronée à 1 unité par défaut (sur x).

2. - Le pouvoir discriminant du test est donné par la valeur $1 - \beta (H)$ donnant la probabilité d'apparition des résultats dans la zone critique déterminée précédemment au cas où $H \neq H_0$ ($p = \frac{1}{2}$) est juste. Au cas où, par exemple, p serait égal à 0,8 et où nous aurions $n = 20$ et où nous utilisons une zone critique bilatérale à seuil $\alpha = 0,05$, nous trouvons dans la table comme valeur critique 5, les zones critiques sont les zones 0 1 2 3 4 5 et 15 16 17 18 19 20.

Au moyen des tables indiquées ou de l'approximation de la loi normale, on pourra calculer la probabilité, dans l'hypothèse où $p = 0,8$, qu'on a de se trouver dans ces zones $1 - \beta (p = 0,8) = 0,80$.

Nous remarquerons en passant que pour une hypothèse H déterminée, le test des signes est "convergent". Ceci signifie que le pouvoir discriminant, dans le cas où l'hypothèse H est juste, a pour limite 1 quand $n \rightarrow \infty$ dans le cas d'une zone bilatérale c'est vrai pour tout $p \neq 1/2$.

dans le cas d'une zone unilatérale droite, c'est vrai pour tout $p > 1/2$

dans le cas d'une zone unilatérale gauche, c'est vrai pour tout $p < 1/2$.

Lorsque n croît, la distance entre les moyennes $\frac{n}{2}$ (H_0 $p = 1/2$) et np (H_0 : $p = p$) soit $n(1/2 - p)$ croît plus vite que les écarts probables $1/2\sqrt{n}$ (H_0) et $\sqrt{np(1-p)}$ (H). La probabilité dans le cas H de se trouver dans la zone critique de H_0 croît donc avec n.

Si n est petit, le test des signes ne conduira à rejeter l'hypothèse $H_0 : p = 1/2$ que si le p réel est voisin de 0 ou 1. Si p est voisin de $1/2$ mais différent, on ne pourra conclure que si l'on augmente n .

Le seuil de confiance α étant fixé, on cherchera à déterminer pour chaque valeur de p la valeur de n nécessaire pour conduire à rejeter l'hypothèse $p_0 = 1/2$ dans r % des cas, ce qui revient donc à se fixer $1 - \beta(p)$, par exemple 95 %.

On aura donc à résoudre les inégalités :

$$2^{-n} \sum_{j=0}^i C_n^j \leq 1/2 \alpha \quad (\text{dans le cas de zone critique bilatérale})$$

et

$$\sum_{j=0}^i C_n^j p^j (1-p)^{n-j} \geq r \quad (\text{dans le cas où } p < 1/2)$$

La table II, à laquelle on arrive ainsi indique les valeurs de n donnant un résultat significatif au moins dans 95 % des cas [$1 - \beta(H_0 : p = p) \geq 0,95$] pour le seuil de confiance α indiqué (zones critiques bilatérales). Si les signes sont distribués dans le rapport 35/65 il faudra ainsi à ce seuil de confiance $\alpha = 0,05$, prendre un échantillonnage d'au moins 143 paires pour obtenir un résultat significatif au moins 95 % du temps ; la valeur critique est de 59. Ces valeurs sont fournies par Mac STEWART "Note on the power of the sign test". [Annals of mathematical statistics Vol 12 (1941) p. 236-238.]

Table II - Valeur minimum de n permettant de déceler les différences significatives dans 95 % des cas pour les diverses proportions réelles p par rapport à l'hypothèse $p_0 = 0,50$; C = valeur critique (zones critiques bilatérales pour : $\alpha = 0,01, 0,05, 0,10, 0,25$).

p	$\alpha = 0,01$		$\alpha = 0,05$		$\alpha = 0,10$		$\alpha = 0,25$	
	N	C	N	C	N	C	N	C
.45(55)	1,777	833	1297	612	1080	512	780	373
.40(60)	442	193	327	145	267	119	193	87
.35(65)	193	78	143	59	118	49	86	37
.30(70)	106	39	79	30	67	26	47	19
.25(75)	66	22	49	17	42	15	32	12
.20(80)	44	13	35	11	28	9	21	7
.15(85)	32	8	23	6	18	5	14	4
.10(90)	24	5	17	4	13	3	11	3
.05(95)	15	2	12	2	11	2	6	1

Les valeurs encadrées sont approchées au maximum à 5 unités près.

3.- Justification des zones critiques employées * (4).

Jusqu'à présent, les zones critiques utilisées ont été les zones critiques classiques et ont compris les queues des distributions. On peut justifier ce choix en appliquant le théorème de J. NEYMAN, "First course in probability and statistics" N.Y. 1950, au rapport de vraisemblance λ (likelihood ratio)

$$\lambda(x; H_0, H_1) = \frac{P[\underline{x} = x / H_0]}{P[\underline{x} = x / H_1]}$$

Si l'on modifie une zone critique pour le test de H_0 en y remplaçant une ou plusieurs valeurs x par des valeurs pour lesquelles $\lambda(x; H_0, H_1)$ est plus grand que le plus grand de ceux des valeurs remplacées, sans que cette modification n'augmente le niveau de confiance α , le pouvoir discriminant du test par rapport à l'hypothèse H_1 diminue :

(4) J. HEMELRIJK - Rapport S95 Centre mathématique d'Amsterdam. Cours de statistique mathématique (1952).

Dans le cas du test des signes

$$\lambda \left(x; \frac{1}{2}, H_1 \right) = \frac{C_n^x 2^{-n}}{C_n^x p^x (1-p)^{n-x}} = \frac{2^{-n}}{p^x (1-p)^{n-x}}$$

Pour n donné :

Si x croît, λ décroît si $p > 1/2$

Si x décroît, λ croît si $p < 1/2$

Ceci signifie que pour tester H_0 contre H_1 ($p > 1/2$), il faut utiliser une zone critique unilatérale droite et pour tester H_0 contre H_1 ($p < 1/2$), il faut une zone critique unilatérale gauche.

Dans le cas où l'on a affaire à un ensemble d'hypothèses, A. WALD "Tests of statistical hypotheses concerning several parameters when the number of observations is large" Am. Math. Soc. Trans. Vol 54 (1943) pp. 426 et suivantes, a démontré que les tests basés sur l'emploi de λ aussi petit que possible possédaient certains propriétés optimum

$$\lambda \left(x; \Omega_0, \Omega \right) = \frac{\text{Max P} \left[\underline{x} = x / H \right]_{H \in \Omega_0}}{\text{Max P} \left[\underline{x} = x / H \right]_{H \in \Omega}}$$

Ω_0 étant l'ensemble des hypothèses à tester contre l'ensemble de toutes les hypothèses Ω .

Si l'on veut tester H_0 ($p_0 = 1/2$) contre les différentes hypothèses $p \neq 1/2$ on a :

$$\lambda = \left(x; \frac{1}{2}, \Omega \right) = \frac{C_n^x 2^{-n}}{\text{Max}_{0 \leq p \leq 1} (C_n^x) p^x q^{n-x}} \quad q = 1 - p$$

pour $x = 0$, le maximum $C_n^x p^x q^{n-x}$ est atteint pour $q = 1$ et $\lambda = 2^{-n}$
 $x = n$ $p = 1$ et $\lambda = 2^{-n}$

$$\lambda = \left(0; \frac{1}{2}, \Omega \right) = \lambda \left(n; \frac{1}{2}, \Omega \right)$$

pour les autres valeurs de x le max : $p^x q^{n-x}$ est obtenu en annulant la dérivée.

$$\begin{aligned} &= x p^{x-1} q^{n-x} - (n-x) p^x q^{n-x-1} \\ &= p^{x-1} q^{n-x-1} [x p - (n-x) q] \\ &= p^{x-1} q^{n-x-1} (x - np) \text{ s'annule pour } p = \frac{x}{n} \end{aligned}$$

la dérivée seconde est $-n \left(\frac{x}{n} \right)^{x-1} \left(1 - \frac{x}{n} \right)^{n-x-1} < 0$, donc on a bien un maximum et

$$\lambda = \frac{2^{-n} n^n}{x^x (n-x)^{n-x}}$$

le dénominateur $x^x (n-x)^{n-x}$ a pour dérivée logarithmique :

$$\frac{d}{dx} \left[x L x + (n-x) L (n-x) \right] = Lx + 1 - L(n-x) - 1 = L \frac{x}{n-x}$$

la dérivée est > 0 pour $x > n-x$ donc $x > \frac{1}{2}n$

$$< 0 \quad x < n-x \quad \text{donc} \quad x < \frac{1}{2}n$$

λ décroît donc si x s'éloigne de $1/2 n$.

Pour $x = 1$ et $x = n-1$

$$\lambda = \frac{2^{-x} n^n}{(n-1)^{n-1}} = 2^{-n} n \cdot \left(\frac{n}{n-1} \right)^{n-1} > 2^{-n}$$

donc supérieur aux valeurs correspondant à $x = 0$ et $x = n$.

Comme λ a la même valeur pour les valeurs de x symétriques par rapport à $\frac{1}{2}n$ on est conduit, pour la recherche des valeurs de x qui rendent λ aussi petit que possible, à utiliser une zone critique bilatérale du type connu.

Si l'on avait à tester Ω_0 ($0 \leq p \leq 1/2$) contre les hypothèses Ω , ($1/2 < p \leq 1$), c'est-à-dire avec Ω : $0 \leq p \leq 1$, on serait conduit par un raisonnement analogue à vérifier le bon emploi de la zone critique unilatérale droite.

4. - Efficacité du test *(3).

Si \underline{z} est une variable distribuée normalement de moyenne a et variance σ^2 , la probabilité d'obtention d'un $z > 0$ est

$$p = \frac{1}{\sqrt{2\pi}} \int_{\frac{0-a}{\sigma}}^{\infty} e^{-1/2 u^2} du.$$

Elle dépend donc de l'estimation $\frac{a}{\sigma}$.

DIXON et MOOD donnent dans leur article comme mesure de l'efficacité celle de l'efficacité de $\frac{a}{\sigma}$. Cette dernière étant la limite du rapport pour $n \rightarrow \infty$ des variances d'une estimation efficace à une estimation donnée par W.G. COCHRAN dans le journal Royal Statistical Society Vol C Part, I (1937) pp. 69-73. "The efficiencies of the binomial series test of significance of a mean and of a correlation coefficient".

La variance de l'estimation efficace de $\frac{a}{\sigma}$:

$$\frac{a}{\sigma} = \frac{\bar{z}}{\sqrt{\frac{\sum (z_i - \bar{z})^2}{n-1}}} = \frac{t}{\sqrt{n}}$$

où t est le t de Student et $\bar{z} = \frac{\sum z_i}{n}$, a la valeur $\frac{1}{n-2}$

La variance donnée par COCHRAN étant, pour n grand, $\frac{2\pi p(1-p)}{n} e^{\left(\frac{a}{\sigma}\right)^2}$

la limite cherchée est alors :

$$\frac{\left(\frac{a}{\sigma}\right)^2}{2\pi p(1-p)}. \text{ Si } \frac{a}{\sigma} = 0 \text{ et } p = \frac{1}{2}$$

l'efficacité est alors :

$$\frac{1}{2\pi \frac{1}{2} \frac{1}{2}} = \frac{2}{\pi} \text{ soit } 63,7 \%$$

Pour n petit, le calcul a été fait et a indiqué au seuil de confiance de 10 % que le test des signes utilisant :

18 paires d'observations était à peu près équivalent au test de Student pour 12 paires	
30	20 à 21 paires
44	28 à 29 paires

Valeurs critiques pour le test des signes,
(zones critiques bilatérales pour le test binomial $p=0,5$)

	Niveau de confiance					Niveau de confiance			
	0,01	0,05	0,10	0,25		0,01	0,05	0,10	0,25
1	-	-	-	-	51	15	18	19	20
2	-	-	-	-	52	16	18	19	21
3	-	-	-	0	53	16	18	20	21
4	-	-	-	0	54	17	19	20	22
5	-	-	0	0	55	17	19	20	22
6	-	0	0	1	56	17	20	21	23
7	-	0	0	1	57	18	20	21	23
8	0	0	1	1	58	18	21	22	24
9	0	1	1	2	59	19	21	22	24
10	0	1	1	2	60	19	21	23	25
11	0	1	2	3	61	20	22	23	25
12	1	2	2	3	62	20	22	24	25
13	1	2	3	3	63	20	23	24	26
14	1	2	3	4	64	21	23	24	26
15	2	3	3	4	65	21	24	25	27
16	2	3	4	5	66	22	24	25	27
17	2	4	4	5	67	22	25	26	28
18	3	4	5	6	68	22	25	26	28
19	3	4	5	6	69	23	25	27	29
20	3	5	5	6	70	23	26	27	29
21	4	5	6	7	71	24	26	28	30
22	4	5	6	7	72	24	27	28	30
23	4	6	7	8	73	25	27	28	31
24	5	6	7	8	74	25	28	29	31
25	5	7	7	9	75	25	28	29	32
26	6	7	8	9	76	26	28	30	32
27	6	7	8	10	77	26	29	30	32
28	6	8	9	10	78	27	29	31	33
29	7	8	9	10	79	27	30	31	33
30	7	9	10	11	80	28	30	32	34
31	7	9	10	11	81	28	31	32	34
32	8	9	10	12	82	28	31	33	35
33	8	10	11	12	83	29	32	33	35
34	9	10	11	13	84	29	32	33	36
35	9	11	12	13	85	30	32	34	36
36	9	11	12	14	86	30	33	34	37
37	10	12	13	14	87	31	33	35	37
38	10	12	13	14	88	31	34	35	38
39	11	12	13	15	89	31	34	36	38
40	11	13	14	15	90	32	35	36	39
41	11	13	14	16	91	32	35	37	39
42	12	14	15	16	92	33	36	37	39
43	12	14	15	17	93	33	36	38	40
44	13	15	16	17	94	34	37	38	40
45	13	15	16	18	95	34	37	38	41
46	13	15	16	18	96	34	37	39	41
47	14	16	17	19	97	35	38	39	42
48	14	16	17	19	98	35	38	40	42
49	15	17	18	19	99	36	39	40	43
50	15	17	18	20	100	36	39	41	43

D'après W.J. DIXON et A.M. MOOD Jrn. of the Am. Stat. Ass. 41 (1946)

Pour les valeurs de $n > 100$, les valeurs critiques approximatives peuvent s'obtenir en prenant les nombres entiers immédiatement inférieurs à

$$\frac{n-1}{2} - k\sqrt{n-1} \text{ où } k = 1,2879 \text{ pour } \alpha = 1\%$$

$$k = 0,9800 \text{ pour } \alpha = 5\%$$

$$k = 0,8224 \text{ pour } \alpha = 10\%$$

$$k = 0,5752 \text{ pour } \alpha = 25\%$$