REVUE DE STATISTIQUE APPLIQUÉE

A. VESSEREAU

Test du nombre de maxima. Caractéristiques et forme limite de la distribution

Revue de statistique appliquée, tome 6, nº 4 (1958), p. 111-117
<http://www.numdam.org/item?id=RSA_1958__6_4_111_0>

TEST DU NOMBRE DE MAXIMA

CARACTÉRISTIQUES ET FORME LIMITE DE LA DISTRIBUTION

par

A. VESSEREAU

Ingénieur en Chef des Manufactures de l'Etat

Dans l'article de M. Meigniez (Revue de Statistique Appliquée-1958-n°11) peu d'indications sont données sur la distribution exacte du "nombre de maxima" dans une suite circulaire de n valeurs. Seule l'espérance mathématique est donnée ($\frac{n}{3}$); c'est empiriquement que la tendance vers la loi normale pour n grand a été constatée.

Dans ce qui suit, on établira, par une méthode utilisant les propriétés des fonctions caractéristiques, les valeurs exactes des quatre premiers moments et des coefficients de Pearson; on en tirera quelques renseignements sur la convergence vers la loi normale; on montrera enfin que la forme limite de la loi est effectivement une loi normale.

On rappelle que k désignant le nombre (aléatoire) de maxima dans une suite de n valeurs, on a, entre les probabilités :

$$P_{k,n} \qquad\qquad P_{k,n-1} \qquad\qquad P_{k-1,\ n-1}$$

la relation de recurrence :

$$\boxed{(1) \qquad (n-1)P_{k,n} = 2k\,P_{k,n-1} + (n-2k+1)P_{k-1,\ n-1}}$$

k varie par valeurs entières de 1 à $\frac{n}{2}$ si n est pair et de 1 à $\frac{n-1}{2}$ si n est impair.

RÉCURRENCE ENTRE FONCTIONS CARACTÉRISTIQUES

On désignera respectivement par $\varphi_n(t)$ et $\varphi_{n-1}(t)$ les fonctions caractéristiques pour des suites de n et (n - 1) valeurs.

La relation (1) peut s'écrire :

$$(n-1)P_{k,n} = 2kP_{k,\ n-1} + (n-1)P_{k-1,\ n-1} - 2\ (k-1)P_{k-1,\ n-1}$$

On supposera n pair, et l'on pourra aisément vérifier que les résultats que l'on va obtenir sont exactement les mêmes lorsque n est impair.

Après avoir multiplié les deux membres de la relation précédente par e^{itk}, on somme de $k = 1$ à $k = \frac{n}{2}$. Il vient :

$$\sum e^{itk} P_{k,n} = \varphi_n(t) \qquad\qquad \sum k e^{itk} P_{k,n-1} = -i \frac{d}{dt} \varphi_{n-1}(t)$$

$$\sum e^{itk} P_{k-1,n-1} = e^{it}\varphi_{n-1}(t) \qquad \sum (k-1)e^{itk}P_{k-1,\ n-1} = -ie^{it}\frac{d}{dt}\varphi_{n-1}(t)$$

D'où la relation de récurrence entre fonctions caractéristiques :

$$\boxed{(2) \qquad \varphi_n(t) = e^{it}\varphi_{n-1}(t) + \frac{2i}{n-1}(e^{it}-1)\frac{d}{dt}\varphi_{n-1}(t)}$$

RÉCURRENCE ENTRE MOMENTS ET CALCUL DES MOMENTS

Les moments pour une suite de n valeurs seront désignés par $\mu'_{1,n}$ (espérance mathématique), $\mu_{2,n}$, $\mu_{3,n}$ ... (moments centrés du 2e, 3e... ordre). Pour une suite de (n - 1) valeurs, on aura de même :

$$\mu'_{1,\ n-1} \qquad \mu_{2,\ n-1} \qquad \mu_{3,\ n-1} \qquad \ldots$$

Les récurrences que l'on obtiendra par la suite seront toutes du type :

$$x_n = \lambda + \frac{n-p}{n-1}\, x_{n-1} \qquad \text{(p entier)}$$

dont la solution (voir démonstration in fine) est :

$$x_n = \frac{\lambda n}{p}$$

1 - RECURRENCE ENTRE LES ESPERANCES MATHEMATIQUES -

En identifiant les termes en (it) dans la relation (2), on obtient :

$$\mu'_{1,n} = 1 + \mu'_{1,n-1} - \frac{2}{n-1}\mu'_{1,n-1}$$

$$\mu'_{1,n} = 1 + \frac{n-3}{n-1}\mu'_{1,n-1}$$

D'où, pour $n \geqslant 3$:

$$\boxed{(3) \qquad \mu'_{1,n} = E_n(k) = \frac{n}{3} \qquad (n \geqslant 3)}$$

2 - RECURRENCE ENTRE LES FONCTIONS CARACTERISTIQUES DES VARIABLES CENTREES -

Pour obtenir plus facilement les récurrences entre moments centrés, on commence par écrire la relation de récurrence entre les fonctions caractéristiques des variables centrées.

En désignant celles-ci par $\psi_n(t)$ et $\psi_{n-1}(t)$, on a :

$$\psi_n(t) = e^{-\frac{nit}{3}}\varphi_n(t) \qquad \psi_{n-1}(t) = e^{-\frac{(n-1)it}{3}}\varphi_{n-1}(t)$$

En portant dans la relation (2), il vient, après quelques simplifications :

$$\boxed{(4) \qquad e^{\frac{it}{3}}\psi_n(t) = \frac{1}{3}(e^{it}+2)\psi_{n-1}(t) + \frac{2i}{n-1}(e^{it}-1)\frac{d}{dt}\psi_{n-1}(t)}$$

On obtient les récurrences entre moments centrés en identifiant dans les deux membres les termes en $\frac{(it)^2}{2!}$, $\frac{(it)^3}{3!}$, ...

3 - RECURRENCE ENTRE LES MOMENTS CENTRES DU 2^e ORDRE (VARIANCES) -

En identifiant les termes en $\frac{(it)^2}{2!}$ on obtient :

$$\mu_{2,n} + \frac{1}{9} = \mu_{2,n-1} + \frac{1}{3} - \frac{4}{n-1}\mu_{2,n-1}$$

$$\mu_{2,n} = \frac{2}{9} + \frac{n-5}{n-1}\mu_{2,n-1}$$

D'où pour $n - 1 \geqslant 3$, c'est-à-dire $n \geqslant 4$:

$$\boxed{(5) \qquad \mu_{2,n} = \sigma_n^2 = \frac{2n}{45} \qquad \sigma_n = \sqrt{\frac{2n}{45}} \qquad (n \geqslant 4)}$$

4 - RECURRENCE ENTRE LES MOMENTS CENTRES DU 3^e ORDRE -

Après identification des termes en $\frac{(it)^3}{3!}$; on obtient :

$$\mu_{3,n} = \frac{n-7}{n-1}\mu_{3,n-1} + \frac{n-7}{n-1}\mu_{2,n-1} - \mu_{2,n} + \frac{8}{27}$$

En remplaçant $\mu_{2,n}$ et $\mu_{2,n-1}$ par les valeurs calculées plus haut, il vient :

$$\mu_{3,n} = -\frac{2}{135} + \frac{n-7}{n-1}\mu_{3,n-1}$$

D'où, pour $n - 1 \geqslant 4$, où $n \geqslant 5$:

$$\boxed{(6) \qquad \mu_{3,n} = -\frac{2n}{945} \qquad (n \geqslant 5)}$$

5 - RECURRENCE ENTRE LES MOMENTS CENTRES DU 4^e ORDRE -

En identifiant les termes en $\frac{(it)^4}{4!}$, on obtient :

$$\mu_{4,n} = \frac{n-9}{n-1}\mu_{4,n-1} + \frac{4(n-10)}{3(n-1)}\mu_{3,n-1} + \frac{2(n-5)}{n-1}\mu_{2,n-1} - \frac{4}{3}\mu_{3,n} - \frac{2}{3}\mu_{2,n} + \frac{26}{81}$$

En remplaçant $\mu_{3,n}$, $\mu_{3,n-1}$, $\mu_{2,n}$, $\mu_{2,n-1}$, par les valeurs calculées plus haut, il vient :

$$\mu_{4,n} = \frac{56n - 90}{945} + \frac{n-9}{n-1}\mu_{4,n-1}$$

On se ramène à une récurrence des types précédents en passant par l'intermédiaire des <u>cumulants</u> :

$$K_{4,n} = \mu_{4,n} - 3\mu_{2,n}^2 = \mu_{4,n} - \frac{12\,n^2}{2025} \qquad K_{4,n-1} = \mu_{4,n-1} - \frac{12\,(n-1)^2}{2025}$$

La relation de récurrence entre les cumulants est :

$$K_{4,n} = -\frac{22}{525} + \frac{n-9}{n-1}K_{4,n-1}$$

D'où pour $n - 1 \geqslant 5$, où $n \geqslant 6$

$$K_{4,n} = -\frac{22\,n}{4725} \qquad \mu_{4,n} = \frac{12\,n^2}{2025} - \frac{22\,n}{4725} = \frac{2n\,(14\,n - 11)}{4725}$$

$$\boxed{(7) \qquad K_{4,n} = -\frac{22\,n}{4725} \qquad \mu_{4,n} = \frac{2\,n\,(14\,n - 11)}{4725} \qquad (n \geqslant 6)}$$

RÉCAPITULATION - CŒFFICIENT DE PEARSON - LOI RÉDUITE

1°) On calcule facilement les coefficients de Pearson :

$$\beta_{1,n} = \frac{\mu_{3,n}^2}{\mu_{2,n}^3} = \frac{5}{98\,n} \qquad \beta_{2,n} = \frac{\mu_{4,n}}{\mu_{2,n}^2} = 3 - \frac{33}{14\,n}$$

2°) Le tableau suivant résume les valeurs des moments et des coefficients de Pearson :

$$\boxed{\begin{array}{llll} & E_n(k) = \frac{n}{3} \quad (n \geqslant 3) & \sigma_n^2 = \frac{2\,n}{45} \quad (n \geqslant 4) & \\ (8) & \mu_{3,n} = -\frac{2\,n}{945} \quad (n \geqslant 5) & \mu_{4,n} = \frac{2\,n\,(14\,n - 11)}{4725} & (n \geqslant 6) \\ & \beta_{1,n} = \frac{5}{98\,n} \quad (n \geqslant 5) & \beta_{2,n} = 3 - \frac{33}{14\,n} \quad (n \geqslant 6) & \end{array}}$$

3°) La variable réduite a pour expression :

$$x = \frac{k - \frac{n}{3}}{\sqrt{2n/45}}$$

$$\boxed{(9) \qquad x = \frac{3\,k - n}{\sqrt{0,4\,n}}}$$

Quand on assimilera la loi de x à une loi continue, on fera la correction pour continuité en écrivant :

$$x = \frac{3\,k - n + 1}{\sqrt{0,4\,n}}$$

4°) Le coefficient $\beta_1 = \frac{5}{98\,n}$ tend très rapidement vers zéro quand n croît: la distribution tend très rapidement à être symétrique.

Le coefficient $\beta_2 = 3 - \frac{33}{14\,n}$ tend vers 3, valeur caractéristique de la loi normale, mais beaucoup plus lentement.

Le tableau ci-dessous donne les valeurs de β_1, et β_2 pour quelques valeurs de n.

n	β_1	β_2
10	0.0051	2.764
20	0.0025	2.882
30	0.0017	2.921
40	0.0013	2.941
50	0.0010	2.953
100	0.0005	2.976
...	......	.....
∞	0.0000	3.000

CONVERGENCE VERS LA LOI NORMALE

Nous admettrons que la loi réduite tend vers une limite lorsque n augmente indéfiniment; on va montrer que cette loi limite est la loi normale réduite.

La relation (4) entre fonctions caractéristiques des variables centrées s'écrit encore :

$$e^{\frac{it}{3}} \psi_{n+1}(t) = \frac{1}{3}(e^{it}+2)\psi_n(t) + \frac{2i}{n}(e^{it}-1)\frac{d}{dt}\psi_n(t)$$

Pour faire apparaître les fonctions caractéristiques $\Phi_n(t)$ et $\Phi_{n+1}(t)$ des variables réduites, on remplace, dans la relation précédente, t par $t = t/\sqrt{\frac{2n}{45}} = \frac{t}{a\sqrt{n}}$ en posant $a = \sqrt{\frac{2}{45}}$.

Il vient :

$$(10)\quad e^{\frac{it}{3a\sqrt{n}}}\psi_{n+1}\left(\frac{t}{a\sqrt{n}}\right) = \frac{1}{3}\left(e^{\frac{it}{a\sqrt{n}}}+2\right)\psi_n\left(\frac{t}{a\sqrt{n}}\right) + \frac{2i}{n}\left(e^{\frac{it}{a\sqrt{n}}}-1\right)(a\sqrt{n})\frac{d}{dt}\psi_n\left(\frac{t}{a\sqrt{n}}\right)$$

On a :

$$\psi_{n+1}\left(\frac{t}{a\sqrt{n}}\right) = \psi_{n+1}\left(\frac{t}{a\sqrt{n+1}}\right) + \frac{t}{a}\left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}}\right)\frac{d}{d\left(\frac{t}{\sqrt{n+1}}\right)}\psi_{n+1}\left(\frac{t}{a\sqrt{n}}\right)$$

$$\text{Comme}\quad \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} = \frac{1}{\sqrt{n}}\left[1 - \sqrt{1-\frac{1}{n+1}}\right] \sim \frac{1}{2n\sqrt{n}}$$

et que : $\psi_{n+1}\left(\frac{t}{a\sqrt{n+1}}\right) = \Phi_{n+1}(t) \qquad \psi_n\left(\frac{t}{a\sqrt{n}}\right) = \Phi_n(t)$

la relation (10) s'écrit :

$$(11)\quad e^{\frac{it}{3a\sqrt{n}}}\left[\Phi_{n+1}(t) + \frac{t}{2n}\frac{d}{dt}\Phi_{n+1}(t)\right] = \frac{1}{3}\left(e^{\frac{it}{a\sqrt{n}}}+2\right)\Phi_n(t) + \frac{2ia}{\sqrt{n}}\left(e^{\frac{it}{a\sqrt{n}}}-1\right)\frac{d}{dt}\Phi_n(t)$$

Pour n très grand $\Phi_{n+1}(t) \sim \Phi_n(t) = \Phi(t)$, la relation (11) s'écrit :

$$\left[e^{\frac{it}{3a\sqrt{n}}} - \frac{1}{3}e^{\frac{it}{a\sqrt{n}}} - \frac{2}{3}\right]\Phi(t) = \left[\frac{2ia}{\sqrt{n}}\left(e^{\frac{it}{a\sqrt{n}}}-1\right) - \frac{t}{2n}e^{\frac{it}{3a\sqrt{n}}}\right]\frac{d\Phi}{dt}$$

On développe :

$$e^{\frac{it}{3a\sqrt{n}}} = 1 + \frac{it}{3a\sqrt{n}} + \frac{1}{2!}\left(\frac{it}{3a\sqrt{n}}\right)^2 + \dots$$

$$e^{\frac{it}{a\sqrt{n}}} = 1 + \frac{it}{a\sqrt{n}} + \frac{1}{2!}\left(\frac{it}{a\sqrt{n}}\right)^2 + \dots$$

et après les simplifications qui apparaissent dans l'exécution des calculs, on trouve :

$$\left[\frac{t}{9\,a^2} + \frac{13\,it^2}{81\,a^2\,\sqrt{n}} + \ldots\right] \Phi(t) = \left[-\frac{5}{2} - \frac{7\,it}{6\,a\sqrt{n}} - \ldots\right] \frac{d\Phi}{dt}$$

équation qui, pour n tendant vers l'infini, se réduit à :

$$\frac{t}{9\,a^2}\,\Phi(t) = -\frac{5}{2}\,\frac{d\Phi}{dt}$$

ou, comme $a^2 = \frac{2}{45}$; $\frac{t}{9\,a^2} = \frac{45\,t}{18} = \frac{5t}{2}$

$$t\Phi(t) = -\frac{d\Phi}{dt}$$

$$\frac{d\Phi}{\Phi} = -\,t\,dt$$

$$\log\Phi = -\frac{t^2}{2} + \lambda \qquad \Phi(t) = A\,e^{-\frac{t^2}{2}}$$

A = 1 puisque l'on doit avoir $\Phi(0) = 1$

$$\boxed{\Phi(t) = e^{-\frac{t^2}{2}}}$$

qui est bien la fonction caractéristique de la loi normale réduite.

NOTE SUR LA RÉCURRENCE

$$x_n = \lambda + \frac{n-p}{n-1}\,x_{n-1}$$

On commence par établir l'expression générale de la somme des produits de k nombres consécutifs.

a) k = 1 $\quad S_1 = 1 + 2 + 3 + \ldots + n$

On a

$$(n+1)n = n(n-1) + 2\,n$$
$$n(n-1) = (n-1)(n-2) + 2\,(n-1)$$

d'où en sommant :

$$n(n+1) = 2\,S_1 \qquad\qquad S_1 = \frac{n(n+1)}{2}$$

b) k = 2 $\quad S_{1.2} = 1.2. + 2.3 + 3.4 + \ldots + (n-1)n$

On a

$$(n+1)n\,(n-1) = n(n-1)(n-2) + 3n(n-1)$$
$$n(n-1)(n-2) = (n-1)\,(n-2)\,(n-3) + 3\,(n-1)\,(n-2)$$

d'où en sommant :

$$(n+1)n\,(n-1) = 3\,S_{1.2} \qquad\qquad S_{1.2} = \frac{(n+1)n\,(n-1)}{3}$$

c) k = 3 $\quad S_{1.2.3.} = 1.2.3 + 2.3.4 + \ldots + (n-2)\,(n-1)n$

On a

$$(n+1)n\,(n-1)\,(n-2) = n(n-1)\,(n-2)\,(n-3) + 4n\,(n-1)\,(n-2)$$

d'où comme précédemment :

$$(n+1)n(n-1)(n-2) = 4S_{1.2.3} \qquad S_{1.2.3} = \frac{(n+1)n(n-1)(n-2)}{4}$$

d) <u>k quelconque</u>. On trouve évidemment :

$$S_{1.2\ldots k} = \frac{(n+1)n(n-1)\ldots(n-k+1)}{k+1}$$

Soit maintenant la récurrence $x_n = \lambda + \frac{n-p}{n-1} x_{n-1}$ (p entier) qui s'écrit encore:

$$(n-1)x_n = \lambda(n-1) + (n-p)x_{n-1}$$

On multiplie par $(n-2)(n-3)\ldots\ldots(n-p+1)$

$$(n-1)(n-2)\ldots(n-p+1)x_n = \lambda(n-1)(n-2)\ldots(n-p+1) + (n-2)(n-3)\ldots(n-p)x_{n-1}$$

$$(n-2)(n-3)\ldots(n-p)x_{n-1} = \lambda(n-2)(n-3)\ldots(n-p) + (n-3)(n-4)\ldots(n-p-1)x_{n-1}$$

d'où en sommant :

$$(n-1)(n-2)\ldots(n-p+1)x_n = \lambda S^{(n-1)}_{1.2\ldots(p-1)} = \lambda\frac{n(n-1)\ldots\ldots(n-p+1)}{p}$$

$$x_n = \frac{\lambda n}{p}$$