

SUR UNE FAMILLE D'INTEGRALES AVEC PARAMÈTRES IMPORTUNS; TESTS DE CONCORDANCE

Fabian Todor

1. Si m juges ont chacun attribué un rang à chaque objet de k objets, alors le coefficient

$$(1.1) \quad W = \frac{S}{S_{\max}} = \frac{12S}{m^2(k^3 - k)}$$

où R_{ij} est le rang de juge i , à l'objet j , $C_j = \sum_{i=1}^m R_{ij}$ et $S = \sum_{j=1}^k [C_j - \bar{C}]^2$ s'appelle le coefficient de concordance de Kendall.

Selon [1] la loi de probabilité de $m(k-1)W$ est approchée par la loi Khi deux avec $k-1$ degrés de liberté. Un coefficient de concordance plus general que W est donné par la formule:

$$(1.2) \quad B = \frac{\sum_{j=1}^k f_j^{\max}}{2km} + \frac{1}{2}W$$

où f_j^{\max} est la fréquence du rang attribué le plus souvent à l'objet j , ou le max des fréquences $\{f_{1j}, f_{2j}, \dots, f_{p_jj}\}$ et où $p_j =$ le nombre des rangs distincts attribués à l'objet j .

2. Selon [1], pour résoudre les problèmes de tests non paramétriques il faut savoir la loi de probabilité pour:

$$(2.3) \quad y = kmB^2.$$

Selon [2] on a le théorème suivant:

"Si les fréquences $\{f_{1j}, f_{2j}, \dots, f_{p_jj}\}$ forment un échantillon sur la variable aléatoire f obéissant à la loi uniforme sur $[0, f_j^{\max}]$ et on note par

$$(2.4) \quad f_{sj} = \text{Sup}\{f_{1j}, f_{2j}, \dots, f_{p_jj}\},$$

alors on a les relations suivantes:

$$(2.5) \quad f_{1j} + f_{2j} + \dots + f_{p_jj} = m$$

et

$$(2.6) \quad E(f_{sj}) = \frac{p_j}{1+p_j} f_j^{\max}.$$

C'est à dire f_{sj} est un estimateur asymptotique non biaisé de f_j^{\max} . Si on note la moyenne arithmétique par :

$$f_{jm} = \frac{f_{1j} + f_{2j} + \dots + f_{p_j j}}{p_j},$$

alors toujours selon [2] on a aussi la relation suivante :

$$(2.7) \quad E(2f_{jm}) = f_j^{\max}.$$

C'est à dire $2f_{jm}$ est un estimateur non biaisé pour f_j^{\max} . Les relations (2.6) et (2.7) donnent la relation suivante :

$$(2.8) \quad E(f_{sj}) = \frac{p_j}{1+p_j} E(2f_{jm}).$$

En éliminant la moyenne E et en remplaçant $f_{jm} p_j = m$, on a :

$$(2.9) \quad f_{sj} = \frac{2m}{1+p_j}.$$

Alors, un estimateur pour la somme $\sum_j f_j^{\max}$ est donné par :

$$(2.10) \quad \sum f_j^{\max} = 2m \sum_j \frac{1}{1+p_j}.$$

Un estimateur pour B de la relation (1,2) aura la forme :

$$(2.11) \quad B_* = \frac{\sum_j \frac{1}{1+p_j}}{k} + \frac{W}{2}$$

On note la variable construite avec p_j par :

$$(2.12) \quad Z_j = \begin{cases} \frac{1}{1+p_j} & \text{si } j=1,2,\dots,k \\ 0 & \text{si non.} \end{cases}$$

La variable Z_j obéit à la loi uniforme de densité suivante :

$$(2.13) \quad f(z) = \begin{cases} \frac{1}{m} & \text{si } Z = Z_j \\ 0 & \text{si non.} \end{cases}$$

La moyenne de la variable Z_j est

$$Z_{jm} = \frac{1}{mk} \sum_{j=1}^k \frac{1}{1+p_j},$$

et alors le premier term de la formule (2.11) sera le produit $Z = m \cdot Z_{jm}$, pendant que l'estimateur de B sera donné par:

$$(2.15) \quad B_* = Z + \frac{W}{2}.$$

Les variables Z et W sont indépendantes par définition de W et par la construction de Z . Selon [3] la densité conjointe de (Z, W) sera donnée par:

$$(2.16) \quad f(z, w) = \frac{e^{-\frac{z^2}{2}}}{(2\pi)^{\frac{1}{2}}} \frac{[m(k-1)]^{(k-1)/2} w^{(k-1)/2} \exp\left\{-\frac{m(k-1)w}{2}\right\}}{2^{(k-1)/2} \Gamma[(k-1)/2]}.$$

Par un changement des variables

$$(2.17) \quad U = m(k-1)W,$$

on a le système:

$$(2.18) \quad W = \frac{U}{m(k-1)} \text{ et } Z = B_* - \frac{U}{2m(k-1)}.$$

Toujours selon [3] le jacobien du système (2.18) sera:

$$J = \begin{vmatrix} 0 & 1/m(k-1) \\ 1 & -\frac{1}{2m(k-1)} \end{vmatrix} = -\frac{1}{m(k-1)},$$

et la densité conjointe de (B_*, U) sera donnée par la formule:

$$(2.19) \quad h(b_*, u) = \frac{\exp\left\{-\frac{1}{2}\left[b_* - \frac{u}{2m(k-1)}\right]^2\right\}}{(2\pi)^{\frac{1}{2}}} \cdot \frac{u^{(k-1)/2} e^{-u/2}}{2^{(k-1)/2} \Gamma[(k-1)/2]}.$$

La densité marginale de B_* sera donnée en integrant par U , et on a:

$$(2.20) \quad h(b_*, \cdot) = \frac{1}{(2\pi)^{\frac{1}{2}} 2^{(k-1)/2} \Gamma[(k-1)/2]} \cdot \int_0^{\infty} e^{-1/2[b_* - u/2m(k-1)]^2} e^{-u/2} u^{(k-1)/2} du.$$

Finalement pour calculer la densité de l'estimateur:

$$B_* = (Y^*/mk)^{\frac{1}{2}},$$

et on a d'après le calcul en utilisant (2.20):

$$(2.21) \quad d(y) \sim d(y^*) = \frac{(y^*)^{-\frac{1}{2}} \cdot e^{-\frac{y^*}{2mk}}}{(2\pi)^{\frac{1}{2}} (mk)^{\frac{1}{2}} 2^{(k-1)/2} \Gamma[(k-1)/2]} \cdot I(y^*)$$

où:

$$(2.22) \quad I(y^*) = \int_0^{\infty} e^{-u/2 - 1/2[u^2/4m^2(k-1)^2 - u(y^*)^{\frac{1}{2}}/m^{\frac{3}{2}}k^{\frac{1}{2}}(k-1)]} u^{(k-1)/2} du.$$

La densité $d(y^*)$ donnée par (2.21) appartient à la famille Khi-deux si on fait l'abstraction de l'intégrale $I(y^*)$ donnée dans (2.22). Considérée isolée, l'intégrale (2.22) contient une fonction de u (variable aléatoire) et une expression en y^* (paramètre importun).

En vain de dériver sous l'intégrale (2.22) par rapport au paramètre y^* pour approfondir l'étude de la densité $d(y) \sim d(y^*)$. Une observation attentive (voire aussi [4]) sur la densité $d(y^*)$ laisse entendre qu'une telle distribution appartient à la famille de densités KOOPMAN-DARMOIS, dont la forme générale de la densité est suivante:

$$(2.23) \quad f(x/\phi) = A(\phi) \cdot B(x) \cdot e^{[\sum H_j(\phi) \cdot G_j(x)]}, \quad \text{où } A \text{ et } H_j, \quad j=1,2,\dots,k$$

sont fonctions arbitraires de paramètre ϕ , tandis que B et G_j , $i=1,2,\dots,k$ sont fonctions arbitraires de x . Le rôle de x pour la densité $d(y^*)$ est joué par u , tandis que le rôle de ϕ est interprété par le paramètre importun y^* .

S'il est connu un échantillon sur la variable x (respective u), on peut construire des Statistiques Suffisantes "conjointes" pour le paramètre ϕ , sous la forme suivante:

$$(2.24) \quad T_j = \sum_j G_j(x_j), \quad j=1,2,\dots,k,$$

respective dans notre cas des Statistique Suffisantes "conjointes" pour y^* .

Le rôle des tests invariants (fonctions d'un invariant maximal), séquentiels est décisif et vient de nous donner nouvelles espoires de progrès dans l'étude de la densité $d(y) \sim d(y^*)$.

REFERENCES

- [1] G. Martineau, "Statistique non paramétrique," Ed. Sci. et Culture Inc., Montréal, 1990.
- [2] P. Jaffard, "Initiation aux méthodes de la statistique et du calcul des probabilités," Masson, Paris, 1976.
- [3] P. I. Bickel and K. A. Doksum, "Mathematical Statistic," Holden-Day, 1977, p. 9.
- [4] M. H. De Groot, "Probability and Statistics," Addison-Wesley Publishing Co., 1986.