Tirage aléatoire simple (plan SI)

Pierre Duchesne

August 1, 2017

Plan aléatoire simple sans remise

- Ce plan est tel que tout échantillon de taille n, où n est fixé à l'avance, possède la même probabilité.
- Ce tirage représente essentiellement ce que l'on a en tête quand on tire dans une urne n boules sans remise qui sont bien mélangées et indétectable au toucher.
- En pratique, il faut des algorithmes (pouvant être plus ou moins compliqués) pour mettre en oeuvre ce plan.

Fonction de probabilité

Étant donné n fixé, chaque échantillon s a donc la probabilité suivante:

$$p(s) = \begin{cases} 1/\binom{N}{n} & \text{si } s \text{ est de taille } n, \\ 0 & \text{sinon.} \end{cases}$$

Toutes les unités ont la même probabilité d'inclusion:

$$p(s \ni k) = \frac{\binom{1}{1}\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N} = \pi_k.$$

Probabilités d'inclusion d'ordre deux

Un raisonnement similaire donne les probabilités d'inclusion d'ordre deux pour deux unités différentes:

$$p(s \ni k, l) = \frac{\binom{2}{2}\binom{N-2}{n-2}}{\binom{N}{n}} = \frac{n(n-1)}{N(N-1)} = \pi_{kl}, k \neq l.$$

Algorithmes pour la mise en oeuvre du tirage SI

On considère deux types d'algorithmes:

- Dans le premier type, lors d'une série de tirages, chaque tirage donne une unité incluse. On poursuit jusqu'à la taille désirée n.
- Dans le second type, pour une série d'expériences, une pour chaque unité dans la base de sondage, selon l'ordre défini dans la base de sondage dont l'analyste a à sa disposition. Pour chaque unité k, l'expérience donne soit le résultat 'unité incluse' ou 'unité non incluse'.

Exemple d'algorithme du premier type

- On sélectionne la première unité en donnant la probabilité $1/N$ à toutes les unités.
- Ne pas remettre l'unité choisie.
- Tirer la seconde unité en donnant une probabilité de $1/(N-1)$ aux unités restantes.
- Ne pas remettre l'unité choisie.
- $\ldots$
- Tirer la *n*ième unité en donnant la probabilité $1/(N-n+1)$ à toutes les unités restantes.
- Fin.

Second exemple d'un algorithme du premier type

- On sélectionne la première unité en donnant la probabilité $1/N$ à toutes les unités.
- Noter l'unité et remettre l'unité.
- Répéter jusqu'à ce que l'on ait obtenu n unités distinctes.
- Il faut ainsi un nombre d'essais valant ν et satisfaisant $\nu \geq n$.
- Le nombre de tirage est aléatoire (ceci peut être considéré un inconvénient).

Premier exemple d'un algorithme du second type (SSW, p. 12)

- Soient $u_1, \ldots, u_N$ des réalisations d'une variable aléatoire $U(0,1)$.
- Si $u_1 < n/N$, alors l'unité $k = 1$ est incluse.
- $\ldots$
- Si $u_k < (n - n_k)/(N - k + 1)$, où n_k est le nombre d'éléments choisis parmi les $k - 1$ éléments dans la liste de la population.
- La procédure arrête lorsque $n_k = n$.

Second exemple d'un algorithme du second type

- Soient $u_1, u_2, \ldots, u_N$ des réalisations d'une variable aléatoire $U(0, 1)$.
- Trier les nombres obtenus par ordre croissant: $u_{(1)} < u_{(2)} < \ldots < u_{(N)}$.
- Les unités obtenues correspondent aux indices des n premiers individus de la liste triée.

Exemple avec *runif*() de R

Considérons une population de taille $N = 10$. On désire un échantillon de taille $n = 3$.

k	*runif*()	k	*runif*()
1	0.001251259	1	0.001251259
2	0.563585314	3	0.193304239
3	0.193304239	7	0.350291452
4	0.808740501	6	0.479873043
5	0.585009308	2	0.563585314
6	0.479873043	5	0.585009308
7	0.350291452	10	0.746604816
8	0.895962401	4	0.808740501
9	0.822840052	9	0.822840052
10	0.746604816	8	0.895962401

L'échantillon de taille $n = 3$ est $\{1, 3, 7\}$.