Solution du problème précédent (n°42)
proposé par Bernard PARZYSZ

Un hebdomadaire organise un concours selon le principe suivant : une question est posée aux lecteurs ; il s'agit pour les participants d'inscrire la réponse sur une carte postale, et de l'envoyer au siège de la revue.
Le règlement précise que le gagnant au concours sera "la personne dont la carte aura été tirée au hasard parmi celles portant une bonne réponse".

Cependant, afin de s'épargner la fastidieuse tâche de trier préalablement les bonnes réponses des mauvaises, les organisateurs décident d'utiliser la procédure suivante : on tire au hasard une carte parmi **toutes** les cartes reçues ; si cette carte indique la bonne réponse, son expéditeur est déclaré gagnant du concours ; sinon, on opère des tirages successifs (sans remise) d'une carte, jusqu'à obtention d'une bonne réponse.

Blaise, qui a envoyé une carte portant la bonne réponse, se demande si cette procédure ne le désavantage pas par rapport à celle qui figure dans le règlement initial.

Qu'en pensez-vous ?

Nous n'avons reçu, pour ce problème, que trois solutions, ayant pour auteurs Pol LE GALL (57 Rombas), Jacques VERDIER (54, Nancy, communication orale) et André VIRICEL (54 Villers-lès-Nancy). Tous trois concluent que ***la probabilité de gain de Blaise est la même dans les deux cas***.

1) A. Viricel, sans formalisation, indique que dans le second cas « *tout se passe comme si les mauvaises réponses n'existaient pas* » ; ce qui, en effet, résulte du fait qu'un tirage donnant une mauvaise carte petit être considéré comme nul et non avenu ; la procédure revient alors à effectuer un tirage dans l'ensemble (les "bonnes" cartes.

Les deux autres solutions formalisent cette idée en se plaçant dans le cadre de la théorie des probabilités : il est toujours réconfortant, en effet, de constater que, sur tel point particulier, le résultat fourni par la théorie n’est pas contraire à l'intuition. Ce qui concourt à montrer que 1° la théorie et 2° la modélisation que l'on a effectuée ne sont pas complètement farfelues.

2) P. Le Gall procède ainsi :
Soient N le nombre de réponses reçues par l’hebdomadaire et n le nombre de réponses exactes. Appelons A l'événement « *Blaise est le gagnant* ».

a) Si l'hebdomadaire respectait la procédure annoncée, on aurait, en tirant au sort parmi les n bonnes réponses, $P(A) = \frac{1}{n}$.

b) Dans la procédure effective, la main innocente chargée du tirage devra, au pire, effectuer $N - n + 1$ tirages, car il y a N - n mauvaises réponses.

Soit A_k l'événement "c'est au $k^{\text{ème}}$ tirage que l'on obtient la première réponse exacte" $(1 \le k \le N - n + 1)$.

Les événements A_k constituent à l'évidence un système complet, et on peut écrire :

$$P(a) = \sum_{k=1}^{N-n+1} P(A \cap A_k) = \sum_{k=1}^{N-n+1} \left(P(A | A_k) \times P(A_k) \right)$$

Or, pour tout k, on a $P(A | A_k) = \frac{1}{n}$ car aucune bonne carte n'a été précédemment tirée (il en reste donc n), et, sachant qu'une bonne carte va être tirée, Blaise a donc une chance sur n d'être gagnant.

On obtient alors $P(a) = \frac{1}{n} \sum_{k=1}^{N-n+1} P(A_k) = \frac{1}{n}$, puisque $\sum_{k=1}^{N-n+1} P(A_k) = 1$ en vertu du fait que les A_k constituent un système complet d'événements.

Donc, en ce qui concerne les chances de succès de Blaise, les deux procédures sont équivalentes.

3) Venons-en maintenant à la solution de J. Verdier qui, quoique paraissant a priori plus calculatoire, ne manque pas non plus d'intérêt (ne serait-ce que parce qu'elle fait intervenir une formule de combinatoire qu'on a rarement l'occasion d'utiliser) :

Nous conservons les notations précédentes, et nous définissons en outre les événements suivants :

F_i « au $i^{\text{ème}}$ tirage, on obtient une réponse incorrecte » ;

B_k « Blaise gagne au $k^{\text{ème}}$ tirage ».

Nous avons bien sûr $P(A) = \sum_{k=1}^{N-n+1} P(B_k)$. Mais en remarquant que $B_k \subset F_1 \cap F_2 \cap \ldots \cap F_{k-1}$, nous pouvons écrire $B_k = F_1 \cap F_2 \ldots \cap F_{k-1} \cap B_k$. La formule des probabilités composées nous donne alors :

$P(B_k) = P(F_1) \times P(F_2 | F_1) \times P(F_3 | F_1 \cap F_2) \times \ldots \times P(B_k | F_1 \cap F_2 \cap \ldots \cap F_{k-1})$, soit :

$$P(B_k) = \frac{N-n}{N} \times \frac{N-n-1}{N-1} \times \frac{N-n-2}{N-2} \times \ldots \times \frac{N-n-k+2}{N-k+2} \times \frac{1}{N-k-1}$$

ou encore : $P(B_k) = \frac{(N-n)!}{(N-n-k+1)!} \times \frac{(N-k)!}{N!} = \frac{(n-1)! C_{N-k}^{n-1}}{n! C_N^n} = \frac{1}{n} \frac{C_{N-k}^{n-1}}{C_N^n}$

D'où : $P(A) = \sum_{k=1}^{N-n+1} P(B_k) = \frac{1}{nC_N^n} \sum_{k=1}^{N-n+1} C_{N-k}^{n-1}$

Évaluons donc $S = \sum_{k=1}^{N-n+1} C_{N-k}^{n-1}$: en posant $\begin{cases} p = n-1 \\ q = N-n \\ j = N-n-k+1 \end{cases}$ il vient

$S = \sum_{j=0}^{N-n} C_{p+j}^{p}$.

Or on sait (ou on démontre par récurrence) que $\sum_{j=0}^{m} C_{p+j}^{p} = C_{p+m+1}^{p-1}$, d'où

$S = C_{p+N-n+1}^{p+1}$. Ou, en revenant aux notations initiales, $S = C_N^n$.

Et, finalement : $P(A) = \frac{1}{nC_N^n} \times C_N^n = \frac{1}{n}$

Conclusion

Je ne sais si les hebdomadaires qui organisent ce type de concours ont fait appel à un "consultant en probabilités" pour savoir si la procédure qui consiste à trier des milliers de cartes, et celle qui consiste à ... ne rien faire, sont équivalentes, mais je crois deviner vers laquelle ils se sont dirigés. Et, si ce sont uniquement des raisons d'économie qui ont guidé ce choix, nous sommes maintenant en mesure de les rassurer : ils pourront continuer a écrire « *Un tirage au sort, effectué parmi toutes les bonnes réponses par Maître X.... huissier, permettra d'attribuer...* » et à faire autre chose. Mais en y réfléchissant bien - me souffle le logicien - il ne s'agit pas d'autre chose, car si l'on tire parmi *toutes* les cartes, on tire *a fortiori* parmi les bonnes...

Le problème du trimestre (n°43)

proposé par Michel THIRY, de NANCY.

On joint deux à deux n points d'un cercle : quel est le nombre **maximal** de régions du disque que définissent ces cordes ?

N.D.L.R. : cet énoncé se trouve dans beaucoup de manuels de lycée pour montrer qu'il faut se méfier des "fausses" relations de récurrence ([1]) ; mais on n'y trouve pas le calcul du nombre de régions en fonction de n.

([1]). 2 régions pour $n = 2$, 4 régions pour $n = 3$, 8 régions pour $n = 4$, 16 régions pour $n = 5$. Peut-on dire "**etc.**" ?