Sur un exercice de probabilité

Par la Commission MOTS.

A propos de l'exercice de probabilité en série D dans le groupement académique de Paris-Créteil-Versailles-Amiens-Lille-Rouen de juin 1991, de nombreuses remarques de collègues sont parvenues à Jean CAPRON, responsable de l'analyse des sujets du baccalauréat, et en particulier une lettre de deux collègues dont nous allons reprendre de longs extraits.

Le présent article a pour but d'aider à réduire, à l'avenir, les ambiguïtés dans les énoncés d'exercices du baccalauréat.

Rappelons l'exercice en question :

> Dans un jeu télévisé, le candidat doit répondre à vingt questions. Pour chacune des questions l'animateur propose au candidat TROIS réponses possibles, UNE SEULE étant la réponse exacte. Les questionnaires sont établis de façon que l'on puisse admettre que :
>
> *i.* Un candidat retenu pour participer au jeu connaît la réponse exacte pour 60% des questions et donne une réponse au hasard pour les autres.
>
> *ii.* Les questions posées lors du jeu sont deux à deux indépendantes.
>
> D'une façon générale, on note $p(A)$ la probabilité de l'événement A. Si A est un événement, $\overline{A}$ désigne l'événement contraire.

1. Soient les événements :
H : «le candidat choisit au hasard la réponse à la première question»
E : «le candidat donne la réponse exacte à la première question».

a : Calculer les probabilités $p(\mathrm{H})$ et $p(\overline{\mathrm{H}})$. En déduire $p(\mathrm{E}\text{ et }\overline{\mathrm{H}})$ (on pourra remarquer que «E ou H» est l'événement certain).
b. Sachant qu'un candidat répond au hasard à la première question, quelle est la probabilité qu'il donne la réponse exacte? En déduire $p(\mathrm{E}\text{ et H})$.

c. Utiliser *a* et *b* pour montrer que $p(\mathrm{E}) = \dfrac{11}{15}$.

N.B. - On donnera les résultats sous forme de fractions irréductibles.

2.*a*. On considère un candidat pris au hasard et on note X la variable aléatoire «Nombre de réponses exactes données par le candidat aux 20 questions du jeu».

Donner la loi de probabilité de X, c'est-à-dire l'expression, en fonction de k (entier compris entre 0 et 20), de la probabilité $p(\mathrm{X} = k)$.

b. Quel est le nombre moyen de bonnes réponses donné par un candidat pris au hasard?

c. Donner une valeur décimale approchée à 10^{-4} près, de la probabilité qu'un candidat pris au hasard donne 20 réponses exactes.

Dans ce texte, on évoque tantôt *un candidat retenu* pour participer au jeu, tantôt *un candidat*, tantôt *un candidat pris au hasard*; cette diversité ne risque-t-elle pas de troubler les élèves?

On suppose que le candidat *connaît la réponse exacte pour 60% des questions*: s'agit-il de 60% des questions qu'on lui pose ou de 60% de l'ensemble des questions susceptibles de lui être posées? Cela est important car ces deux interprétations conduisent pour la deuxième question à des résultats différents.

On dit que les questions posées sont *deux à deux indépendantes*: comment doit-on l'interpréter? Cela signifie-t-il
- que la connaissance de la réponse exacte à une question n'entraîne pas forcément la connaissance de la réponse exacte à l'autre question?
- ou bien que les deux questions sont posées indépendamment l'une de l'autre (ce qui impliquerait qu'on peut poser deux fois la même question)?

Dans la première question (*) on demande à l'élève de *calculer* $p(\mathrm{H})$ et $p(\overline{\mathrm{H}})$: or $p(\overline{\mathrm{H}})$ est donné, il n'y a donc pas à le calculer ; là aussi, l'élève risque d'être troublé. On lui demande ensuite d'en *déduire* $p(\mathrm{E}\text{ et }\overline{\mathrm{H}})$ en lui faisant remarquer que «E ou H» est l'événement certain : là encore, il y a de quoi être perturbé car $p(\mathrm{E}\text{ et }\overline{\mathrm{H}})$ ne se déduit pas de $p(\mathrm{H})$ et $p(\overline{\mathrm{H}})$, mais s'obtient en postulant implicitement que «$\overline{\mathrm{E}}$ et $\overline{\mathrm{H}}$» est l'événement impossible, ou encore que $p(\mathrm{E}/\overline{\mathrm{H}}) = 1$.

Dans la deuxième question, la formulation du *a* est incorrecte: X étant une fonction, il serait préférable de parler de la variable aléatoire X «qui à chaque jeu associe le nombre de réponses exactes données par le candidat aux 20 questions du jeu». C'est ici qu'il y a deux interprétations possibles qui conduisent à des résultats différents.

Première interprétation: les questionnaires sont préétablis; ils comportent 20 questions dont 60% (soit 12 questions) ont une réponse connue de tout candidat retenu ; le nombre de réponses exactes données par le candidat est supérieur ou égal à 12.

Si $0 \leq k < 12$, alors $p(X = k) = 0$.

Si $12 \leq k \leq 20$, le candidat se retrouve dans la situation d'un questionnaire à 8 questions auxquelles il lui faut répondre au hasard ; il donne à chaque question la réponse exacte avec une probabilité de 1/3 ou une réponse fausse avec une probabilité de 2/3; d'où $p\,(X = k) = \mathrm{C}_8^{k-12}\left(\frac{1}{3}\right)^{k-12} \times \left(\frac{2}{3}\right)^{8-(k-12)}$.

Deuxième interprétation: les questionnaires comportent 20 questions; ces questions ont été tirées au hasard dans un recueil de questions, ce recueil ayant été établi de telle sorte que tout candidat retenu connaisse 60% des réponses aux questions du recueil. On est donc en présence d'une situation de 20 épreuves répétées et le schéma de Bernoulli ne s'applique que si le résultat à une épreuve ne dépend pas des résultats précédents. Or, il s'agit de tirages successifs de 20 questions, mais *sans remise* (on n'a jamais vu poser deux fois la même question à un candidat!). Il est clair que si le candidat a déjà été confronté à 15 questions dont il connaît la réponse, la probabilité pour qu'il connaisse la réponse à la 16ème question n'est plus de 60%, mais est strictement inférieur à 60%. On ne peut calculer cette probabilité que si l'on connaît le nombre de questions du recueil; mais si celui-ci est «assez important», on peut supposer qu'elle est pratiquement égale à 60% ; encore fallait-il le dire explicitement: «On admet que le nombre de questions dont on dis-

(*) L'expression «choisit au hasard» est antinomique.

pose est suffisamment grand pour que la situation du candidat face à une question soit indépendante du fait qu'il connaisse ou non la réponse aux questions précédentes». Dans ce cas, bien entendu :

si $0 \leq k \leq 20$, alors $p\,(X = k) = C_{20}^{k}\left(\dfrac{11}{15}\right)^{k} \times \left(\dfrac{4}{15}\right)^{20-k}$ ce que vraisemblablement on attendait des élèves!

Dans le *b* de la deuxième question,, on demande *le nombre moyen de bonnes réponses donné par un candidat pris au hasard*. Il s'agit vraisemblablement du nombre moyen de réponses donné*es* par un candidat retenu. Dans ce cas, pourquoi ne pas avoir tout simplement demandé à l'élève de calculer l'espérance mathématique E(X) de la variable aléatoire X ?

Rappelons, pour conclure, que dans l'analyse des sujets du baccalauréat 1991, Jean CAPRON n'a reçu aucune critique concernant les autres sujets donnés en probabilité, ce qui dénote un progrès par rapport aux années précédentes.