Exercice| [4371] | 1| Utiliser l'inégalité de Bienaymé-Tchebychev

On sait que 40% des individus d'une population possèdent un certain caractère C. On considère un échantillon de 200 personnes de cette population.
Peut-on affirmer, à plus de 80%, que dans cet échantillon il y a entre 30 et 50% de personnes ayant le caractère C ?

Pistes de réflexion

- Il s'agit d'introduire ici une variable aléatoire à laquelle on appliquera l'inégalité de Bienaymé-Tchebychev.
- On pensera notamment à transformer le problème en un problème de prélèvement dans une population.

Éléments de correction

On met en place le protocole suivant :

- on effectue un prélèvement de 200 individus dans une population dont on suppose la taille suffisamment grande pour pouvoir assimiler ce prélèvement à des prélèvements successifs et avec remise, de sorte que la proportion p d'individus de cette population possédant le caractère C puisse être considérée comme quasiment constante égale à $0,4$.
- on désigne par X la variable aléatoire égale au nombre d'individus du prélèvement de taille 200 qui possèdent le caractère C.
- l'hypothèse faite sur la taille de la population permet de considérer que l'on réalise donc successivement 200 fois de manière indépendante la même épreuve de Bernoulli qui consiste à prélever avec un remise un individu de la population et de regarder s'il possède ou non le caractère C, qu'il aura avec la probabilité $p = 0,4$.

La variable aléatoire X ainsi définie suit alors une loi binomiale de paramètres $n = 200$ et $p = 0,4$.
On aura en particulier $\mathbb{E}(X) = 80$ et $\mathbb{V}(X) = 48$.
La fréquence d'apparition du caractère C dans cet échantillon est donc égal à $\frac{X}{200}$, et on cherche à estimer la valeur de l'événement $\left[0,3 \leq \frac{X}{200} \leq 0,5\right]$.

Or on a :
$$\begin{aligned}
\left[0,3 \leq \frac{X}{200} \leq 0,5\right] &= [60 \leq X \leq 100] \\
&= [-20 \leq X - 80 \leq 20] \\
&= [-20 \leq X - \mathbb{E}(X) \leq 20] \\
&= [|X - \mathbb{E}(X)| \leq 20] \\
&= \overline{[|X - \mathbb{E}(x)| > 20]} \\
&\underset{|X-\mathbb{E}(X)|\in\mathbb{N}}{=} \overline{[|X - \mathbb{E}(x)| \geq 21]}
\end{aligned}$$

Or d'après l'inégalité de Bienaymé-Tchebychev appliqué à la variable aléatoire X avec $\varepsilon = 21$, on a :
$$\mathbb{P}([|X - \mathbb{E}(X)| \geq 21]) \leq \frac{28}{441}$$

Ainsi, il vient que :
$$\begin{aligned}
& \left(\mathbb{P}([|X - \mathbb{E}(X)| \geq 21]) \leq \frac{48}{441}\right) \\
\Leftrightarrow\ & \left(1 - \mathbb{P}\left(\overline{[|X - \mathbb{E}(X)| \geq 21]}\right) \leq \frac{28}{441}\right) \\
\Leftrightarrow\ & \left(1 - \mathbb{P}([|X - \mathbb{E}(X)| \leq 20]) \leq \frac{48}{441}\right) \\
\Leftrightarrow\ & \left(\mathbb{P}([|X - \mathbb{E}(X)| \leq 20]) \geq 1 - \frac{48}{441}\right) \\
\Leftrightarrow\ & \left(\mathbb{P}([|X - \mathbb{E}(X)| \leq 20]) \geq \frac{393}{441}\right) \\
\Leftrightarrow\ & \left(\mathbb{P}([0,3 \leq X \leq 0,5]) \geq \frac{393}{441}\right)
\end{aligned}$$

Comme $\frac{393}{441} \approx 0,89$, on est sûr à plus de 89%, qu'il y a entre 30 et 50% de personnes de cet échantillon qui possèdent le caractère C.