Concentration, loi des grands nombres

L'objectif de cette section est d'une part d'approfondir le sens de l'écart-type comme mesure de dispersion, d'autre part de couronner la partie « Probabilités » par la loi des grands nombres, qui est le premier résultat fondamental de la théorie des probabilités et dont les implications sont considérables.

Pour cela, l'outil employé est l'inégalité de Bienaymé-Tchebychev dont l'idée fondamentale est mise en valeur : l'écart type σ d'une variable aléatoire X est l'unité naturelle pour étudier la dispersion de X autour de son espérance μ ; par construction, il est naturel d'observer des écarts de X à μ en deçà ou au-delà de σ. L'inégalité de Bienaymé-Tchebychev montre qu'en revanche des écarts de X à μ de quelques σ deviennent improbables. Ce résultat, d'une importance majeure en lui-même, permet de plus d'établir la loi des grands nombres, selon laquelle l'écart entre la moyenne d'un échantillon d'une variable aléatoire et l'espérance de cette variable ne dépasse une valeur donnée à l'avance qu'avec une probabilité qui tend vers zéro quand la taille de l'échantillon tend vers l'infini.

Il est utile de faire remarquer aux élèves que le caractère universel de l'inégalité de Bienaymé-Tchebychev a pour contrepartie le fait qu'elle est loin d'être optimale : ainsi, elle montre qu'un écart à μ supérieur à 2σ est de probabilité inférieure ou égale à 1/4 alors que les élèves ont découvert par simulation que cette probabilité est souvent majorée par 0,05. En avoir conscience ne diminue pas l'intérêt théorique de l'inégalité de Bienaymé-Tchebychev, et permet de mettre en évidence des cas de raisonnement par conditions suffisantes, par exemple la recherche d'une taille d'échantillon pour majorer une probabilité.

Contenus

- Inégalité de Bienaymé-Tchebychev. Pour une variable aléatoire X d'espérance μ et de variance V, et quel que soit le réel strictement positif δ : $P(|X-\mu| \geqslant \delta) \leqslant \dfrac{V(X)}{\delta^2}$.
- Inégalité de concentration. Si M_n est la variable aléatoire moyenne d'un échantillon de taille n d'une variable aléatoire d'espérance μ et de variance V, alors pour tout $\delta > 0$, $P(|M_n-\mu| \geqslant \delta) \leqslant \dfrac{V}{n\delta^2}$
- Loi des grands nombres.

Capacité attendue

- Appliquer l'inégalité de Bienaymé-Tchebychev pour définir une taille d'échantillon, en fonction de la précision et du risque choisi.

Exemples d'algorithme

- Calculer la probabilité de $(|S_n - pn| > \sqrt{n})$, où S_n est une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n,p)$. Comparer avec l'inégalité de Bienaymé-Tchebychev.
- Simulation d'une marche aléatoire.
- Simuler N échantillons de taille n d'une variable aléatoire d'espérance μ et d'écart type σ. Calculer l'écart type s de la série des moyennes des échantillons observés, à comparer à $\sigma/\sqrt{n}$. Calculer la proportion des échantillons pour lesquels l'écart entre la moyenne et μ est inférieur ou égal à ks, ou à $k\sigma/\sqrt{n}$, pour $k = 1, 2, 3$.

Approfondissements possibles

- Estimation.
- Marche aléatoire.
- Exemples d'application issus d'autres disciplines pour diverses valeurs de n : sondage (par exemple $n = 1\,000$), étude du sex ratio (par exemple $n = 10^6$), demi-vie d'atomes radioactifs ($n = 10^{23}$).