3. Probabilités et statistique

On approfondit le travail en probabilités et statistique mené les années précédentes.
Afin de traiter les champs de problèmes associés aux données continues, on introduit les lois de probabilité à densité. Le programme en propose quelques exemples et, en particulier, la loi normale qui permet notamment d'initier les élèves à la statistique inférentielle par la détermination d'un intervalle de confiance pour une proportion à un niveau de confiance de 95 %.
Cette partie se prête particulièrement à l'étude de problèmes issus d'autres disciplines.
Le recours aux représentations graphiques et aux simulations est indispensable.

Contenus	Capacités attendues	Commentaires
Conditionnement, indépendance		
Conditionnement par un événement de probabilité non nulle. Notation $P_A(B)$.	• Construire un arbre pondéré en lien avec une situation donnée. • Exploiter la lecture d'un arbre pondéré pour déterminer des probabilités. • Calculer la probabilité d'un événement connaissant ses probabilités conditionnelles relatives à une partition de l'univers.	On représente une situation à l'aide d'un arbre pondéré ou d'un tableau. On énonce et on justifie les règles de construction et d'utilisation des arbres pondérés. Un arbre pondéré correctement construit constitue une preuve. Le vocabulaire lié à la formule des probabilités totales n'est pas un attendu du programme, mais la mise en œuvre de cette formule doit être maîtrisée.
Indépendance de deux événements.	▣ Démontrer que si deux événements A et B sont indépendants, alors il en est de même pour $\overline{A}$ et B.	Cette partie du programme se prête particulièrement à l'étude de situations concrètes. ◇ Des activités algorithmiques sont menées dans ce cadre, notamment pour simuler une marche aléatoire. ⇆ [SVT] Hérédité, génétique, risque génétique.
Notion de loi à densité à partir d'exemples Loi à densité sur un intervalle.		Les exemples étudiés s'appuient sur une expérience aléatoire et un univers associé Ω, muni d'une probabilité. On définit alors une variable aléatoire X, fonction de Ω dans **R**, qui associe à chaque issue un nombre réel d'un intervalle I de **R**. On admet que X satisfait aux conditions qui permettent de définir la probabilité de l'événement $\{X \in J\}$ comme aire du domaine : $\{M(x, y)\,;\, x \in J \text{ et } 0 \leq y \leq f(x)\}$ où f désigne la fonction de densité de la loi et J un intervalle inclus dans I. Toute théorie générale des lois à densité et des intégrales sur un intervalle non borné est exclue.

Contenus	Capacités attendues	Commentaires
Loi uniforme sur $[a,b]$. Espérance d'une variable aléatoire suivant une loi uniforme.	• Connaître la fonction de densité de la loi uniforme sur $[a,b]$.	L'instruction « nombre aléatoire » d'un logiciel ou d'une calculatrice permet d'introduire la loi uniforme sur $[0,1]$. La notion d'espérance d'une variable aléatoire à densité f sur $[a,b]$ est introduite à cette occasion par $\mathrm{E}(X)=\int_a^b t\,f(t)\,\mathrm{d}t$. On note que cette définition constitue un prolongement dans le cadre continu de l'espérance d'une variable aléatoire discrète. (AP) *Méthode de Monte-Carlo.*
Lois exponentielles.	• Calculer une probabilité dans le cadre d'une loi exponentielle.	▣ On démontre qu'une variable aléatoire T suivant une loi exponentielle vérifie la propriété de durée de vie sans vieillissement : pour tous réels t et h positifs, $P_{T\geq t}(T\geq t+h)=P(T\geq h)$.
Espérance d'une variable aléatoire suivant une loi exponentielle.	▣ Démontrer que l'espérance d'une variable aléatoire suivant une loi exponentielle de paramètre λ est $\frac{1}{\lambda}$.	L'espérance est définie comme la limite quand x tend vers $+\infty$ de $\int_0^x t\,f(t)\,\mathrm{d}t$ où f est la fonction de densité de la loi exponentielle considérée. Cette partie du programme se prête particulièrement à l'étude de situations concrètes, par exemple sur la radioactivité ou la durée de fonctionnement d'un système non soumis à un phénomène d'usure.

Contenus	Capacités attendues	Commentaires
Loi normale centrée réduite $\mathcal{N}(0,1)$. Théorème de Moivre Laplace (admis).	• Connaître la fonction de densité de la loi normale $\mathcal{N}(0,1)$ et sa représentation graphique. ◙ Démontrer que pour $\alpha \in \left]0,1\right[$, il existe un unique réel positif u_α tel que $P(-u_\alpha \leq X \leq u_\alpha) = 1-\alpha$ lorsque X suit la loi normale $\mathcal{N}(0,1)$. • Connaître les valeurs approchées $u_{0,05} \approx 1{,}96$ et $u_{0,01} \approx 2{,}58$.	Pour introduire la loi normale $\mathcal{N}(0,1)$, on s'appuie sur l'observation des représentations graphiques de la loi de la variable aléatoire $Z_n = \frac{X_n - np}{\sqrt{np(1-p)}}$ où X_n suit la loi binomiale $\mathcal{B}(n, p)$, et cela pour de grandes valeurs de n et une valeur de p fixée entre 0 et 1. Le théorème de Moivre Laplace assure que pour tous réels a et b, $P(Z_n \in [a,b])$ tend vers $\int_a^b \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{x^2}{2}} \mathrm{d}x$ lorsque n tend vers $+\infty$. L'espérance d'une variable aléatoire suivant la loi $\mathcal{N}(0,1)$ est définie par $\lim_{x \to -\infty} \int_x^0 t\, f(t)\mathrm{d}t + \lim_{y \to +\infty} \int_0^y t\, f(t)\mathrm{d}t$ où f désigne la densité de cette loi. On peut établir qu'elle vaut 0. On admet que la variance, définie par $\mathrm{E}\left((X - \mathrm{E}(X))^2\right)$, vaut 1.
Loi normale $\mathcal{N}(\mu, \sigma^2)$ d'espérance μ et d'écart-type σ.	• Utiliser une calculatrice ou un tableur pour calculer une probabilité dans le cadre d'une loi normale $\mathcal{N}(\mu, \sigma^2)$. • Connaître une valeur approchée de la probabilité des événements suivants : $\{X \in [\mu - \sigma, \mu + \sigma]\}$, $\{X \in [\mu - 2\sigma, \mu + 2\sigma]\}$ et $\{X \in [\mu - 3\sigma, \mu + 3\sigma]\}$, lorsque X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$.	Une variable aléatoire X suit une loi $\mathcal{N}(\mu, \sigma^2)$ si $\frac{X-\mu}{\sigma}$ suit la loi normale $\mathcal{N}(0,1)$. On fait percevoir l'information apportée par la valeur de l'écart-type. ⇆ [SI et SPC] Mesures physiques sur un système réel en essai. La connaissance d'une expression algébrique de la fonction de densité de la loi $\mathcal{N}(\mu, \sigma^2)$ n'est pas un attendu du programme. On illustre ces nouvelles notions par des exemples issus des autres disciplines.

<table>
<tr><th>Contenus</th><th>Capacités attendues</th><th>Commentaires</th></tr>
<tr><td>Intervalle de fluctuation</td><td>◙ Démontrer que si la variable aléatoire X_n suit la loi $\mathcal{B}(n, p)$, alors, pour tout α dans $]0,1[$ on a,
$\lim\limits_{n\to+\infty} P\left(\frac{X_n}{n} \in I_n\right) = 1-\alpha$,
où I_n désigne l'intervalle
$\left[p - u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + u_\alpha\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$.

• Connaître l'intervalle de fluctuation asymptotique(*) au seuil de 95 % :
$\left[p - 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}, p + 1,96\frac{\sqrt{p(1-p)}}{\sqrt{n}}\right]$ où p désigne la proportion dans la population.</td><td>La démonstration ci-contre donne l'expression d'un intervalle de fluctuation asymptotique(*) au seuil $1-\alpha$ de la variable aléatoire fréquence $F_n = \frac{X_n}{n}$ qui, à tout échantillon de taille n, associe la fréquence obtenue f.

Avec les exigences usuelles de précision, on pratique cette approximation dès que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$.

En majorant $1,96\sqrt{p(1-p)}$, on retrouve l'intervalle de fluctuation présenté en classe de seconde.

La problématique de prise de décision, déjà rencontrée, est travaillée à nouveau avec l'intervalle de fluctuation asymptotique.</td></tr>
<tr><td>Estimation

Intervalle de confiance (*).

Niveau de confiance.</td><td>• Estimer par intervalle une proportion inconnue à partir d'un échantillon.

• Déterminer une taille d'échantillon suffisante pour obtenir, avec une précision donnée, une estimation d'une proportion au niveau de confiance 0,95.</td><td>Les attendus de ce paragraphe sont modestes et sont à exploiter en lien avec les autres disciplines.

◙ Il est intéressant de démontrer que, pour une valeur de p fixée, l'intervalle $\left[F_n - \frac{1}{\sqrt{n}}, F_n + \frac{1}{\sqrt{n}}\right]$ contient, pour n assez grand, la proportion p avec une probabilité au moins égale à 0,95.

On énonce alors que p est élément de l'intervalle $\left[f - \frac{1}{\sqrt{n}}, f + \frac{1}{\sqrt{n}}\right]$ avec un niveau de confiance de plus de 95 %, où f désigne la fréquence observée sur un échantillon de taille n.
Avec les exigences usuelles de précision, on utilise cet intervalle dès que $n \geq 30$, $np \geq 5$ et $n(1-p) \geq 5$.

La simulation de sondages sur tableur permet de sensibiliser aux fourchettes de sondage.</td></tr>
</table>

		Il est important de noter que, dans d'autres champs, on utilise l'intervalle $\left[f-1{,}96\frac{\sqrt{f(1-f)}}{\sqrt{n}}, f+1{,}96\frac{\sqrt{f(1-f)}}{\sqrt{n}}\right]$ qu'il n'est pas possible de justifier dans ce programme. ⇆ [SVT] Analyse de graphiques où les données sont fournies par des intervalles de confiance. AP *Prise de décision lors de la comparaison de deux proportions (par exemple lors d'un essai thérapeutique).*

(*)Avec les notations précédentes :

Un intervalle de fluctuation asymptotique de la variable aléatoire F_n au seuil $1-\alpha$ est un intervalle déterminé à partir de p et de n et qui contient F_n avec une probabilité d'autant plus proche de $1-\alpha$ que n est grand.

Un intervalle de confiance pour une proportion p à un niveau de confiance $1-\alpha$ est la réalisation, à partir d'un échantillon, d'un intervalle aléatoire contenant la proportion p avec une probabilité supérieure ou égale à $1-\alpha$, intervalle aléatoire déterminé à partir de la variable aléatoire fréquence F_n qui, à tout échantillon de taille n, associe la fréquence.
Les intervalles de confiance considérés ici sont centrés en la fréquence observée f .

Algorithmique

En seconde, les élèves ont conçu et mis en œuvre quelques algorithmes. Cette formation se poursuit tout au long du cycle terminal.

Dans le cadre de cette activité algorithmique, les élèves sont entraînés à :
- décrire certains algorithmes en langage naturel ou dans un langage symbolique ;
- en réaliser quelques-uns à l'aide d'un tableur ou d'un programme sur calculatrice ou avec un logiciel adapté ;
- interpréter des algorithmes plus complexes.

Aucun langage, aucun logiciel n'est imposé.

L'algorithmique a une place naturelle dans tous les champs des mathématiques et les problèmes posés doivent être en relation avec les autres parties du programme (analyse, géométrie, statistiques et probabilités, logique) mais aussi avec les autres disciplines ou le traitement de problèmes concrets.
À l'occasion de l'écriture d'algorithmes et de programmes, il convient de donner aux élèves de bonnes habitudes de rigueur et de les entraîner aux pratiques systématiques de vérification et de contrôle.

Instructions élémentaires (affectation, calcul, entrée, sortie) Les élèves, dans le cadre d'une résolution de problèmes, doivent être capables : • d'écrire une formule permettant un calcul ; • d'écrire un programme calculant et donnant la valeur d'une fonction, ainsi que les instructions d'entrées et sorties nécessaires au traitement.
Boucle et itérateur, instruction conditionnelle Les élèves, dans le cadre d'une résolution de problèmes, doivent être capables de : • programmer un calcul itératif, le nombre d'itérations étant donné ; • programmer une instruction conditionnelle, un calcul itératif, avec une fin de boucle conditionnelle.