**Titre:** Connexions entre Espaces de Travail : une étude entre probabilités et algorithmique

**Auteurs:** LAVAL Dominique et TRUNKENWALD Jannick

**Publication:** Actes du huitième colloque de l'Espace Mathématique Francophone – EMF 2022

**Directeur:** Adolphe Cossi Adihou, Université de Sherbrooke (Canada/Bénin) avec l'appui des membres du comité scientifique et des responsables des groupes de travail et projets spéciaux

**Éditeur:** Les Éditions de l'Université de Sherbrooke

**Année:** 2023

**Pages:** 446 - 462

**ISBN:** 978-2-7622-0366-0

**URI:**

**DOI:**

# Connexions entre Espaces de Travail : une étude entre probabilités et algorithmique

LAVAL Dominique[1] – TRUNKENWALD Jannick[2]

**Résumé** – En 10e année (classe de 2nde), l'apprenant découvre la fluctuation de l'échantillonnage. Notre attention se porte sur l'apport des algorithmes à la compréhension de l'intervalle de prédiction dans le domaine des probabilités. Une approche algorithmique de simulation d'échantillonnage est proposée aux enseignants du Lycée Français d'Alger (Algérie) dans le cadre d'un travail de formation. Ce travail est analysé dans sa genèse sémiotique, discursive et instrumentale, au sein des différents champs concernés.

**Mots-clefs :** algorithmique, espace de travail, simulation, fréquence, modélisation.

**Abstract** – In grade 10, learner discovers sampling fluctuation. Our attention focuses on the contribution of algorithms to the understanding of prediction interval in the field of probabilities. An algorithmic approach of sampling simulation is proposed to teachers at the French High School of Algiers (Algeria) as part of a training work. The work is analyzed in its semiotic, discursive and instrumental genesis, within the different fields involved.

**Keywords:** algorithmic, workspace, simulation, frequency, modelling.

1. CY Cergy Paris Université, INSPE Académie de Versailles, France, dominique.laval@cyu.fr

2. Université Paris Cité, LDAR, France, jannick.trunkenwald@yahoo.fr

## Introduction

Bien que ce ne soit pas la partie la plus abstraite du programme de mathématiques en 2nde, les statistiques, depuis plus de trois décennies s'introduisent partout et il est indispensable que les élèves et leurs enseignants connaissent les mécanismes de base qui se cachent derrière les nombres dont nous sommes quotidiennement abreuvés. Dès la 2nde, la notion d'intervalle de confiance est abordée. Néanmoins, nous observons l'absence d'un référentiel théorique conséquent dans le domaine des probabilités qui soit à disposition des élèves. Nous faisons alors l'hypothèse qu'une utilisation de l'outil informatique dans le cadre d'une approche expérimentale va permettre une autre forme de validation des notions abordées. Nous situant dans la continuité des travaux de Trunkenwald et Laval (2019) sur une formation visant l'introduction en classe d'une simulation algorithmique d'un lancer de dés, une nouvelle expérimentation est proposée auprès de professeurs de mathématiques du lycée français d'Alger. Nous faisons le choix d'exposer ici une situation dans le domaine des probabilités associée à une situation géométrique à laquelle les enseignants sont confrontés. Seules des considérations d'ordre didactique guident notre réflexion lors du choix de la tâche, ainsi que sur les analyses faites du travail mathématique et algorithmique en jeu.

## Une tâche pour aborder l'approche fréquentiste

### *Un objectif*

Selon le document « Ressources pour la classe de seconde, Probabilités et statistiques » (2009) :

> […] On peut, par expérimentation et simulation, faire observer aux élèves que les échantillons de taille *n* obtenus à partir d'un modèle de Bernoulli ont, pour environ 95 % d'entre eux, des fréquences d'apparition du nombre 1 qui fluctuent dans un intervalle centré en *p* et d'amplitude $\frac{2}{\sqrt{n}}$

Nous précisons au passage que les bornes de cet intervalle peuvent être rigoureusement déterminées à l'aide d'une table décrivant une loi binomiale de paramètres *n* et *p*. Cependant, cette approche probabiliste, liée aussi au domaine combinatoire, ne constitue pas l'objet de la présente réflexion. Par conséquent, elle ne sera pas décrite dans cette communication. Ainsi, cette prise de conscience de la variabilité naturelle d'un phénomène, due au seul hasard, peut permettre la méthode statistique en aidant l'apprenant à définir la « confiance » à accorder aux estimations. Une approche « fréquentiste » à travers des observations de fréquences relatives de caractères représentatifs de ces phénomènes, peut favoriser la compréhension de la notion de fréquence théorique d'un événement (Laval, 2018). Parzysz (2009) considère que l'introduction de l'intervalle de confiance au niveau de la 2nde se fait sans justification formelle.

> « L’accès à la notion de modèle […] peut être préparé par l’étude et la simulation d’expériences aléatoires diverses correspondants au même modèle probabiliste. La comparaison des procédures, des tableaux et des hypothèses sous-jacentes doit permettre aux élèves de se convaincre de l’analogie que présentent ces expériences sous leurs apparences diverses, et de déboucher sur la notion de schéma d’expérience, constituant en quelque sorte une classe d’équivalence d’expériences aléatoires. » Parzysz (2009, p. 102)

Par ailleurs, Laval (2018, p. 448) observe que « […] les probabilités dans le secondaire, associées aux statistiques, constituent un domaine où des situations extra-mathématiques jouent un rôle fondamental pour l’introduction de l’aléatoire. » Ainsi, la simulation sur ordinateur de phénomènes aléatoires est un moyen utile pour « obtenir des séries statistiques de taille suffisante pour observer la convergence des fréquences relatives et des moyennes statistiques » (Ibid.). Selon Laval (2018), la simulation d’un phénomène aléatoire suppose l’utilisation d’une modélisation de la situation. Dans la continuité du cycle de modélisation de Blum et Leiss (2005), complété par Nechache (2016) pour un modèle probabiliste de type numérique, nous adaptons celui-ci à l’apport des spécificités de l’algorithmique[3] (Figure 1).

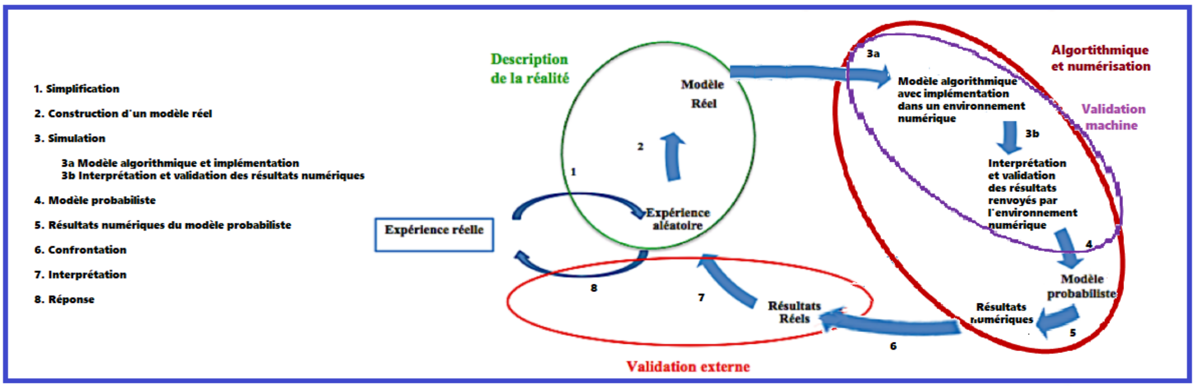



**Fig. 1 :** Cycle de modélisation avec un modèle probabiliste de type algorithmique

À la suite des travaux de Parzysz (2009) et de Nechache (2016), Trunkenwald (2018) explicite l’interaction entre les domaines statistique et probabiliste lorsque l’expérience réelle est simplifiée sous forme d’une expérience aléatoire permettant d’envisager un premier modèle probabiliste réel. Celui-ci est alors traduit en un premier modèle algorithmique associé à la simulation pour passer ensuite à un modèle probabiliste numérique. L’implémentation et l’exécution de l’algorithme de simulation donne des résultats numériques que doit interpréter l’utilisateur. Ces résultats sont d’abord vus par rapport au modèle réel, puis interprétés en regard de l’expérience aléatoire. Cette dernière peut être ajustée suivant l’information apportée par l’expérience réelle. Trunkenwald observe qu’« un nouveau modèle réel peut aussi être envisagé pour affiner la compréhension de l’expérience réelle (ce qui

3. L’algorithmique est la science des algorithmes. Un algorithme est une suite finie et ordonnée d’instructions successives qui manipulent diverses données pour réaliser une tâche précise par son utilisateur. Par ailleurs, lorsque l’on définit un algorithme, celui-ci ne doit contenir que des instructions compréhensibles par celui qui devra l’exécuter (machine ou humain).

correspond alors à un nouveau cycle) ». Par ailleurs, tenant compte de Lagrange et Laval (2019) sur les connexions entre espaces de travail, nous faisons l'hypothèse que l'introduction de l'algorithmique et d'un environnement numérique dans le processus de modélisation va aider l'apprenant à surmonter le coût en nombre de manipulations et permettre une démarche empirique satisfaisante. Mais exploiter cet environnement informatique nécessite d'enclencher un nouveau processus de modélisation spécifique. Nous souhaitons aborder cette problématique spécifique de réalisation d'algorithmes de simulation qui, une fois implémentés dans un environnement numérique, serait conjointe à une exploitation d'un point de vue probabiliste des résultats obtenus. Nous formulons l'hypothèse que les connexions entre modèle algorithmique et modèle probabiliste final simulant la fluctuation d'échantillonnage peuvent être élaborés dans de meilleures conditions si on fait subir, au modèle algorithmique, plusieurs cycles de modélisations au sens de Trunkenwald (2018). La nature de ces cycles fait l'objet de cet article.

### *Vers le choix d'une tâche et son expérimentation avec les premières analyses*

Le programme de 2$^{nde}$ conduit à la découverte de la fluctuation d'échantillonnage avec la propriété suivante : « Soit $p$ la probabilité de succès, et soit $n$ le nombre d'expériences aléatoires répétées. Toute fréquence $f$ de succès vérifie $f \in \left[p - \frac{1}{\sqrt{n}} ; p + \frac{1}{\sqrt{n}}\right]$ avec un niveau de confiance d'au moins 95 % ». A ce niveau, le processus de découverte de cet intervalle reste expérimental et basé sur l'observation d'échantillons qui ont la même taille. Chaque valeur de fréquence correspond à un échantillon obtenu lui-même en répétant l'expérience aléatoire considérée. Il s'agit d'une approche empirique basée sur un raisonnement inductif. De plus, le « coût en temps » correspondant à la nécessité de répéter la même expérience un grand nombre de fois entraîne un besoin d'automatisation. Cette idée est facilitée par l'utilisation d'un calcul basé sur une opération de générateur pseudo-aléatoire. Par ailleurs, l'approche fréquentiste peut nécessiter de considérer plusieurs domaines, tant mathématiques qu'informatique. En effet, un environnement numérique associé à l'implémentation d'un algorithme permet de réduire le temps passé à répéter l'expérience aléatoire. Cependant, des compétences spécifiques sont nécessaires au domaine algorithmique : reconnaissance des variables informatiques, des instructions conditionnelles et des structures de boucles, syntaxe du langage machine, interface utilisateur, etc., et le manque d'expertise en informatique peut entraîner des difficultés supplémentaires pour les apprenants. Cela peut également causer une préoccupation métacognitive (voire cognitive) aux enseignants. L'utilisation du champ numérique permet d'automatiser un protocole expérimental qui nécessite l'utilisation de gestes, d'une syntaxe et des modes de pensée spécifiques. Nous pouvons aussi nous interroger sur l'utilisation d'une méthode algorithmique implémentable dans un environnement numérique, pour une meilleure compréhension de l'approche fréquentiste, d'un point de vue mathématique. Nous pensons que l'écriture d'algorithmes de simulation correspond à une étape de formalisation plus élevée pour les différents protocoles statistiques impliqués, et que le rôle du « langage » prend une importance particulière dans ce sens. Une formation sur la simulation de fluctuation d'échantillonnage en 2$^{nde}$ est réalisée lors d'un stage au Lycée Français d'Alger en 2018. Cinq groupes (A, B, C, D et E) de trois enseignants doivent résoudre le problème :

| |
|---|
| Quelle est la probabilité que la distance entre deux points choisis au hasard le long d'un S-segment de droite, soit supérieure à la moitié de la longueur du S-segment ? |

Nous considérons cette tâche comme représentative des attentes institutionnelles d'une approche fréquentiste. En effet, il est possible d'explorer la fluctuation de l'échantillonnage avec une répétition réelle de l'expérience aléatoire, comme les segmentations de ficelles aux ciseaux, mais compte tenu de la temporalité de ce processus aléatoire lors d'une approche manuelle de la situation, nous faisons l'hypothèse que les enseignants proposent un modèle algorithmique de simulation avec un générateur pseudo-aléatoire de type Alea() fournissant un nombre aléatoire situé entre 0 et 1 simulant mathématiquement le hasard. Ici, Alea() va retourner un nombre aléatoire entre 0 et 1 simulant la position d'un point choisi au hasard sur le S-segment. Le travail de modélisation, mettant en évidence les fluctuations d'échantillonnage, peut trouver ainsi son expression formalisée, à travers une simulation algorithmique. L'expérience aléatoire fournissant la distance entre deux points sélectionnés peut être simulée par un calcul de type |Alea() – Alea()|. Nous considérons une telle instruction de calcul comme un algorithme de niveau A0. Il est appelé à chaque simulation de l'expérience aléatoire. Cependant, pour une plus grande automatisation, nous construisons un second algorithme, de niveau A1, à partir de l'algorithme précédent où nous avons une répétition automatisée par la machine de l'expérience aléatoire pour obtenir directement une fréquence de réussite simulée (Cf. figure 3a des annexes). Par ailleurs, le fait d'exécuter plusieurs fois l'algorithme de niveau A1 permet de générer une liste de fréquences, qui peut être représentée sous la forme d'un nuage de points. A cet effet, une troisième étape est envisagée dans le domaine algorithmique avec la construction et l'implémentation en machine d'un algorithme de niveau A2 (Cf. les algorithmes des niveaux A1 et A2 écrits en langage pseudo-code des figures 3a et 3b des annexes) permettant l'automatisation de l'agencement du nuage de points. L'algorithme de niveau A2 est construit à partir de l'algorithme de niveau A1, en répétant la fréquence de réussite simulée pour tracer directement le nuage de points.

Dans le cadre de notre observation, nous faisons le choix que les groupes de travail n'aient accès à aucun fichier préexistant pour accomplir la tâche. Nous observons que les groupes parviennent plus ou moins facilement, avec l'aide du formateur quand cela s'avère nécessaire, à réaliser un modèle algorithmique pour la simulation d'un calcul de fréquence. Ne parvenant pas à construire un algorithme en langage pseudo-code, nous observons que le groupe E fait le choix d'utiliser alors le tableur. Par ailleurs, nous constatons que seul le groupe C parvient à finaliser l'algorithme de niveau A2. Le passage du niveau A1 au niveau A2 correspond à une étape délicate, en particulier une écriture algorithmique d'un *protocole statistique* qui consiste à exploiter les résultats fournis par l'algorithme de niveau A1. Nous faisons l'hypothèse qu'il s'agit d'un obstacle épistémologique lié à l'encapsulation d'un algorithme dans un autre, au sens d'un sous-programme, afin d'automatiser l'exécution du premier. D'autres difficultés aux connections des champs des statistiques et de l'algorithmique sont identifiées au cours de la formation (un extrait des échanges entre le formateur et le groupe B, sur l'introduction d'un compteur dans l'algorithme de niveau A1 est présentée ci-après (Figure 2)).

| | |
|---|---|
| Stagiaire 1 : | Le compteur moi je veux qu'il compte. Donc c'est la position… |
| Tuteur : | D'accord. Donc comment on peut faire un compteur ? |
| Stagiaire 1 : | C'est… Donc il faut l'initialiser… |
| Tuteur : | D'accord. L'initialiser à combien ? |
| Stagiaire 1 : | Au départ de 0, puisqu'il n'y a aucun… mais… |
| Tuteur : | Et quand est-ce que vous l'augmentez ? |
| Stagiaire 1 : | C'est ça ! C'est la position où le mettre !... Est-ce que… |
| Tuteur : | Il va compter quoi ? |
| Stagiaire 2 : | Il va compter le nombre de cas favorables… |
| Tuteur : | … Voilà, donc à côté du « oui » vous l'augmentez de 1. |
| Stagiaire 1 : | … avant ou après ? … |
| Tuteur : | Avant ou après. Au même niveau. |
| Stagiaire 2 : | On l'a mis ici… ça n'a pas marché. |
| Tuteur : | Vous mettez compteur devient compteur + 1. |
| Stagiaire 1 : | On l'a mis, mais ça n'a pas… |

| | |
|---|---|
| Tuteur : | Où est votre compteur ? |
| Stagiaire 1 : | n. |
| Tuteur : | Ah non, n ce n'est pas le compteur, c'est le numéro de l'étape. |
| Stagiaire 1 : | Oui c'est ça ! |
| Tuteur : | Vous devez ajouter une autre variable qui, elle, va compter les succès. n c'est autre chose… n c'est le numéro de votre étape parmi les 100 étapes. |
| Stagiaire 1 : | Donc il nous faut deux… Donc on aura deux compteurs. |
| Tuteur : | Il y a un compteur de l'étape en cours et un compteur du nombre de succès. |
| Stagiaire 1 : | Voilà ! Mais le compteur du nombre de succès c'est à la fin de la boucle ? Et après le sinon… |
| Tuteur : | Il prend sa valeur à la fin de la boucle. Quand vous êtes sortis de la boucle. |
| Stagiaire 1 : | Voilà donc c'est après le « sinon ». |
| Tuteur : | La fin de la boucle ce n'est pas dans le « sinon ». La fin de la boucle c'est après le « Fin pour ». |

**Fig. 2 :** Echanges entre le formateur et des enseignants du groupe B

Dans cet extrait, nous constatons que la prise en compte d'un compteur avec son initialisation, son incrémentation conditionnée et son exploitation finale apparaît comme une construction complexe, pour certains stagiaires. Sa rédaction est intégrée dans trois niveaux du script de l'algorithme, en tenant compte des autres instructions. Ces stagiaires essaient de repérer une position précise qui n'existe pas vraiment pour le compteur. Nous observons ainsi une confusion entre « compteur de réussite » et « compteur de pas de simulation ». Le « compteur de pas » est un processus naturel en statistique qui devient une difficulté sur le plan algorithmique. Enfin, une difficulté similaire apparaît lors d'une confusion entre les consignes « valeur résultat » fournies par le compteur, et « incrémentation » des valeurs intermédiaires utilisées pour son calcul. Nous observons également une confusion entre le champ des probabilités et celui des statistiques. L'enseignant (stagiaire 2) parle de « cas favorable » au lieu de « succès obtenu ». De plus, les transcriptions de la session montrent des difficultés qui sont aussi liées à la modélisation de l'expérience aléatoire. Nous pensons qu'une analyse plus poussée des travaux dans les domaines mathématiques et de l'algorithmique avec une étude de leurs connexions, peut nous aider à affiner notre compréhension de certaines difficultés rencontrées par les enseignants (stagiaires).

## Le choix d'un cadre théorique et la question de recherche

Nous faisons l'hypothèse que les activités impliquant une tâche mathématique associée à plusieurs domaines peuvent être analysées en considérant que les entités impliquées dans la tâche apparaissent sous différentes représentations sémiotiques, instrumentales et discursives. Duval (1993) nomme « registres » les représentations organisées en systèmes sémiotiques. Des processus de manipulations et de conversions se situent à l'intérieur et entre ces registres, et cette description est également complétée par la notion d'« instruments » dans l'activité mathématique de l'apprenant. Dans une « approche instrumental », Rabardel (1995) décrit l'utilisation des technologies numériques pour enseigner et apprendre les mathématiques dans différents contextes. Un tel cadre composé d'« objets » issus des champs mathématiques avec des relations entre eux, et différentes expressions ou images mentales associées, a été abordé par Douady (1986). Les savoirs instrumentaux sont indissociables des savoirs mathématiques. Dans le cadre des *Espaces de Travail* (Kuzniak, 2011), nous souhaitons prendre en compte les relations induites par les signes instrumentaux et mathématiques. Nous émettons l'hypothèse que la coordination de l'activité dans plusieurs domaines peut apporter des réponses dans ce sens. En effet, en algorithmique, l'activité peut se situer dans une approche instrumentale de résolution de la tâche mathématique, avec un système de signes spécifiques à l'informatique. En revanche, dans le domaine des mathématiques, l'activité consiste à résoudre la tâche d'un point de vue mathématique et à utiliser des signes permettant de comprendre et de justifier. Nous souhaitons alors choisir un cadre théorique approprié qui peut donner du sens au travail de l'apprenant lors d'une activité impliquant plusieurs champs avec leurs connexions, compte tenu des genèses sémiotiques, instrumentales et discursives propres à chaque champ. Nous supposons que les articulations entre les *Espaces de Travail Mathématique Spécifique* ($ETM_s$) (Kuzniak & Richard, 2014) et les *Espaces de Travail Algorithmique* (ETA) (Laval, 2018), peuvent permettre d'étudier les différentes genèses des representamen[4] et des processus, du point de vue des $ETM_S$ et ETA idoines des apprenants, mais aussi d'aborder les interactions entre les probabilités/statistiques et l'algorithmique. Par ailleurs, de nombreux travaux de recherche récents et en cours, réalisés dans divers champs mathématiques (analyse, théorie élémentaire des nombres, probabilités discrètes) et informatique sur les modèles des $ETM_s$/ETA, ont permis de prendre conscience des connexions possibles entre les *Espaces de Travail* (Laval, 2018 ; Lagrange & Laval, 2019). Ci-après, nous présentons brièvement ce cadre théorique des *Espaces de Travail*. La question de recherche consiste alors, à savoir comment ce cadre permet d'analyser le travail d'un apprenant impliquant une activité algorithmique et de programmation dans une approche fréquentiste de l'intervalle de fluctuation.

## Un cadre théorique – Les premières analyses

### *Les Espaces de Travails de référence*

Dans la continuité des travaux de Laval (2021), « nous faisons l'hypothèse de ne pas réduire l'algorithmique à un simple sous-domaine des mathématiques ». En effet, nous voyons l'algorithmique comme une discipline à part entière qui peut être en interaction avec des $ETM_s$. Par conséquent, il s'avère de construire et de justifier l'existence d'un ETA comme *Espace de Travail* à part entière. Par ailleurs, afin d'étudier plus finement le travail des enseignants (stagiaires) dans des champs issus des mathématiques et dans un domaine non nécessairement mathématique (l'algorithmique est vue ici comme une discipline à part entière) où cependant, les objets mathématiques peuvent avoir du sens, nous nous référons aux *Espaces de Travail Connectés* (ETC) (Lagrange & Laval, 2019), « en tenant toujours compte de manière exhaustive des dimensions [genèses] sémiotique, discursive et instrumentale » (Laval, 2021). Pour finir, nous nous questionnons sur l'utilité des ETC en lien avec des $ETM_{analyse}$, $ETM_{statistique}$ et des ETA pour relever le défi posé par des activités algorithmiques et de programmation lors de l'introduction des intervalles de fluctuation en Seconde. Par ailleurs, la tâche considérée ici implique plusieurs champs disciplinaires. Pour chacun de ces champs, nous avons un *Espace de Travail*. Ainsi, le cadre des ETC est introduit afin de rendre compte de la façon dont les

4. Au sens sémiotique, les representamen sont des objet ayant la possibilité de signifier.

connexions entre les *Espaces de Travail* donnent du sens aux concepts impliqués. Ce que nous attendons de ce cadre est d'aider à construire et à analyser des situations sur un sujet donné impliquant un domaine mathématique et un connexe, en identifiant les trois genèses dans les *Espaces de Travail* correspondants, en les contrastant et en recherchant des possibilités de connexion. Le cadre des ETM permet de caractériser la façon dont les concepts vont prendre du sens dans un contexte de travail donné. Selon Kuzniak & Richard (2014), le travail dans un ETM s'organise autour de trois genèses : (GS) *Sémiotique*, avec une utilisation de symboles, de graphismes, d'objets concrets compris comme des signes ; (GI) *Instrumental*, avec des constructions faites à partir d'artefacts, comme tableaux, graphes, programme de calculs, arbres, etc. ; (GD) *Discursif*, où se situent les aspects de preuve et de justification avec l'utilisation d'un référentiel théorique. Ces trois genèses (GS, GI et GD) génèrent chacune une relation dynamique entre un plan épistémologique et un plan cognitif (Cf. figure 1 des annexes). Les outils technologiques (artefacts), sémiotiques (representamen) et théoriques (référentiels) du plan épistémologique peuvent être exploités à l'aide de schémas appropriés. Dans ce cas, on se réfère respectivement à une genèse instrumentale, sémiotique et discursive, qui produisent respectivement des construits, des visualisations et des preuves dans le plan cognitif. L'activation de deux genèses peut entraîner une circulation entre les genèses qui les portent, ce qui peut parfois être assimilé à une idée de modélisation (Laval, 2018). Nous appelons « projection » des ETM ou ETA (Trunkenwald & Laval, 2019), la restriction à un champ spécifique de ce que nous analysons dans l'*Espace de Travail* considéré. Nous écrivons respectivement $\text{ETM}_{\text{probabilité}}$ et $\text{ETM}_{\text{statistique}}$, les « projections » dans les ETM sur les champs de probabilités et de statistiques (Cf. figure 2 des annexes).

## *Les premières analyses*

Afin d'éviter une surcharge visuelle, nous faisons le choix de présenter des schémas de ces « projections » avec une vue de dessus (Cf. figures 2, 3a et 3b des annexes). Les points et les traits gras représentent les différentes genèses discursives (GD), instrumentales (GI) et sémiotiques (GS), ainsi que les circulations dans les plans situés entre les axes de ces genèses. La principale connexion entre ETM et ETA est liée à l'utilisation d'un générateur pseudo-aléatoire. Cet artefact est d'abord instrumentalisé au sein d'un ETA pour construire un modèle algorithmique d'une simulation de l'expérience aléatoire. Le lien entre cette construction et la formulation associée provoque une circulation « Instrumentale – Discursive » (I-D) et débouche sur une genèse discursive enrichissant le référentiel ETA. Selon Laval (2018, 2021), les travaux de conception algorithmique présentent une triple dimension instrumentale, sémiotique et discursive au sein des ETA. Nous associons cette dimension discursive à l'idée de « langage ». En effet, les algorithmes de niveaux A0, A1 et A2 sont des écritures formalisées de protocoles expérimentaux. Les expressions algorithmiques en langage naturel, avec leurs trois structures encapsulées, constituent alors une formulation explicite de la marche à suivre dans le domaine statistique pour effectuer une approche fréquentiste de l'intervalle de fluctuation. Nous faisons l'hypothèse que cette explication peut être associée à l'idée de preuve. Cette considération nous ramène à la genèse discursive qui, selon Laval (2018), se confirme lorsque le programme est capable de tourner (à un premier niveau de paradigme correspondant à une vision intuitive de l'algorithme).

Bien que nous ayons fait le choix de ne pas présenter une analyse de la tâche autour d'une utilisation du tableur, nous notons que la dimension discursive des ETA a tendance à s'estomper dans le plan « Sémiotico – Instrumental » (S-I), lorsqu'on travaille avec un tableur. En effet, dans le cas du tableur, les aspects syntaxiques sont plus liés à la nature des artefacts, même si la dimension discursive est encore activée dans le plan cognitif pour mobiliser les outils du référentiel dans l'ETA. Une construction complexe avec un tableur peut aussi réduire la « visualisation cognitive » d'un phénomène et donc la genèse sémiotique, pour se limiter à une genèse instrumentale. Enfin, l'ETA peut perdre toute sa pertinence lorsque l'algorithme implémenté dans un environnement numérique est fourni aux stagiaires pour être exécuté et utilisé sans aucun travail de modification attendu dans sa propre structure. En effet, dans ce cas, l'algorithme se trouve réduit à un artefact prêt à être instrumentalisé en « projection » d'un ETM sur un autre champ. Les « programmes de simulation » sont très actifs dans la dimension instrumentale des ETM, en dehors des ETA. Ils sont instrumentalisés comme des outils pour construire une observation statistique, au sein d'$\text{ETM}_{\text{statistique}}$ et au sein d'$\text{ETM}_{\text{probabilité}}$, si la tâche présente des propriétés issues du hasard.

### *Les analyses : cas des algorithmes A0, A1 et A2 écrits en pseudo-code*

L'algorithme de niveau A2 vise à reprendre le protocole expérimental autour du dessin d'un nuage de points. Nous attendons des enseignants (stagiaires) qu'ils automatisent le nuage de points des fréquences en améliorant l'algorithme de niveau A1 donnant la simulation d'une fréquence. Ensuite, la mise en œuvre de l'algorithme de niveau A2 permet d'observer différents nuages, et de conjecturer la propriété de l'intervalle de fluctuation. Nous proposons une série de figures (Cf. figures 3a et 3b en annexes) illustrant le schéma éclaté (vue de dessus des ETMs et ETA) correspondant aux conceptions et exécutions de chaque niveau des algorithmes. Plus précisément, en ce qui concerne les interactions entre probabilités et statistiques, une description plus détaillée est présentée par Trunkenwald (2018) pour la tâche choisie. Au niveau A0, nous observons que dans l'ETA, l'algorithme est conçu dans une genèse instrumentale de l'artefact du générateur pseudo-aléatoire. Celui-ci simule une distance entre deux points choisis au hasard le long du S-segment. Cette modélisation algorithmique génère une validation discursive de la simulation expérimentale aléatoire, à travers une genèse discursive dans l'ETA. Pour rappel, l'algorithme de niveau A0 est ensuite utilisé comme outil technologique dans l'$\text{ETM}_{\text{statistique}}$ à travers une genèse instrumentale dans ce même $\text{ETM}_{\text{statistique}}$. On active ainsi la dimension discursive en mobilisant la fréquence, et la dimension sémiotique en mobilisant un type de gestion des données. Un protocole expérimental de calcul d'une fréquence est alors construit en exécutant plusieurs fois l'algorithme de niveau A0. Au niveau A1 (Cf. figure 3a des annexes), le protocole de simulation de fréquences est totalement automatisé. Ici, le travail algorithmique génère une circulation complète dans l'ETA. L'algorithme de niveau A1, correspondant à un nouvel algorithme de simulation de fréquences, est ensuite utilisé comme outil technologique dans l'$\text{ETM}_{\text{statistique}}$ à travers une genèse instrumentale, avec une activation de la dimension discursive en mobilisant un système de coordonnées cartésiennes. Ce nouveau protocole permet de construire un nuage de points des fréquences, en exécutant plusieurs fois, de manière automatisée, l'algorithme de niveau

A1. Nous observons que ce travail génère une genèse sémiotique dans l'$ETM_{statistique}$ en visualisant les fluctuations de l'échantillon. Nous appuyant sur cette nouvelle connexion de type sémiotique entre les concepts de fréquence et de probabilité, nous observons que dans l'$ETM_{statistique}$, la genèse sémiotique débouche sur une genèse sémiotique dans l'$ETM_{probabilité}$. Nous interprétons cela comme une « visualisation cognitive » des fluctuations du hasard. Au niveau A2 (Cf. figure 3b des annexes), le protocole de construction du nuage de points des fréquences est ici complétement automatisé. Le travail algorithmique génère à nouveau une circulation complète dans l'ETA. Cet algorithme de niveau A2, qui permet de construire le nuage de points (Cf. figure 3b des annexes), est ensuite utilisé comme outil technologique dans l'$ETM_{statistique}$ à travers une genèse instrumentale. Cette construction d'un nouveau protocole conduit ainsi à une circulation complète dans l'$ETM_{statistique}$ permettant de construire une conjecture des expressions des bornes de l'intervalle de fluctuation à l'aide d'un raisonnement inductif. Ce travail génère une genèse discursive dans l'$ETM_{statistique}$ avec cette nouvelle compréhension du rôle de l'intervalle, à la lumière des fluctuations de l'échantillon. Appuyée sur ce nouveau lien discursif entre fréquence et probabilité, la genèse discursive dans l'$ETM_{statistique}$ débouche sur une genèse discursive dans l'$ETM_{probabilité}$, avec une compréhension du sens des formules des bornes de l'intervalle de fluctuation, à la lumière de l'approche fréquentiste de la probabilité. Dans le cas de l'algorithme de niveau A2, nous pouvons compléter l'analyse faite précédemment en faisant le lien entre le travail dans les ETM/ETA et le cycle de modélisation modifié (Cf. figure 4 des annexes).

## Conclusion

Les trois phases algorithmiques (A0, A1 et A2) décrites précédemment permettent une construction par encapsulations successives de l'algorithme de simulation. Ces trois étapes algorithmiques entrent également en résonance avec les phases expérimentales du protocole statistique. Cette congruence sémantique trouve son expression dans un certain aspect du travail algorithmique : une organisation de procédures de type « sous-programmes » en blocs. Dans l'$ETM_{statistique}$, nous mettons en évidence les connexions entre les trois domaines considérés : l'algorithme A0 conçu en connexion entre la dimension instrumentale des $ETM_s$ et ETA. Ensuite, cet algorithme A0 est exécuté afin de construire un nouveau protocole expérimental. L'algorithme A0 est ainsi considéré comme un outil technologique à travers une genèse instrumentale dans l'$ETM_{statistique}$. Puis cet algorithme, pris comme outil théorique dans une genèse discursive de l'ETA, est utilisé comme un nouvel objet algorithmique visant à automatiser un nouveau protocole expérimental, à travers l'algorithme A1. Deux nouveaux cycles identiques entre $ETM_{statistique}$ et ETA connectés peuvent alors être générés successivement. Nous construisons ainsi l'algorithme A2, qui permet d'établir une conjecture pour encadrer les fluctuations. Ensuite, cette conjecture permet de visualiser dans une genèse sémiotique de l'$ETM_{probabilité}$ les bornes de l'intervalle de fluctuation, afin d'activer la genèse discursive de l'$ETM_{probabilité}$. Nous enrichissons ainsi le référentiel théorique sur la propriété de l'intervalle de fluctuation. Par ailleurs, afin de mettre en évidence la dialectique des cycles des connexions évolutives entre les trois domaines (probabilités, statistiques et algorithmique), notre analyse peut être utilisée pour concevoir un prototype de formation des enseignants de lycée, mais aussi une activité pour des élèves de 2[nde]. L'objectif serait d'éviter de confronter les apprenants aux difficultés simultanées dans les différents domaines concernés. De même, un tel prototype peut également être utilisé comme outil de remédiation.

## Annexes

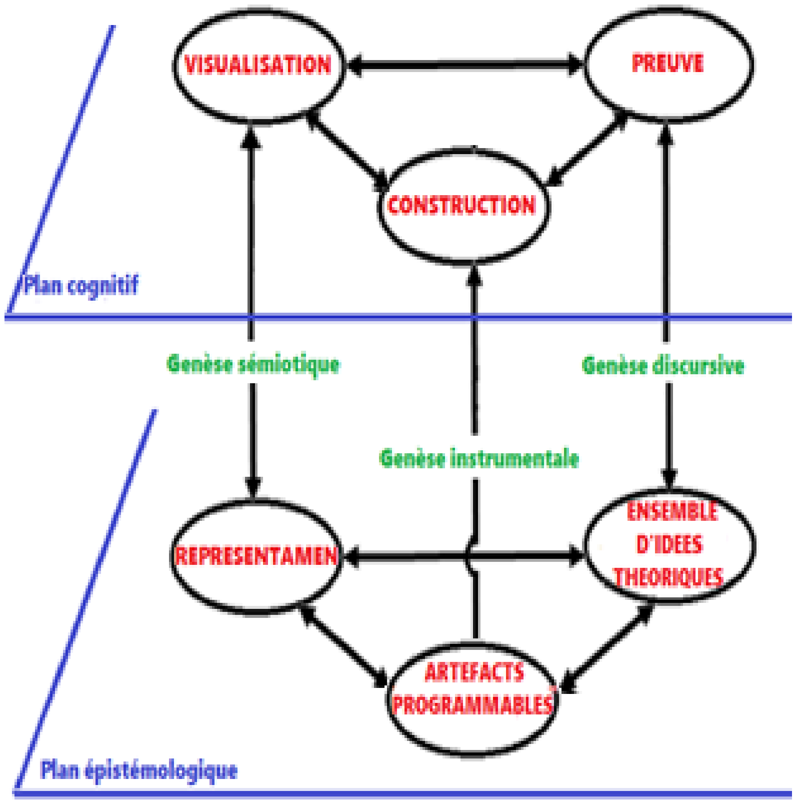



**Fig. 1 :** Représentations des ETA (Laval, 2018)

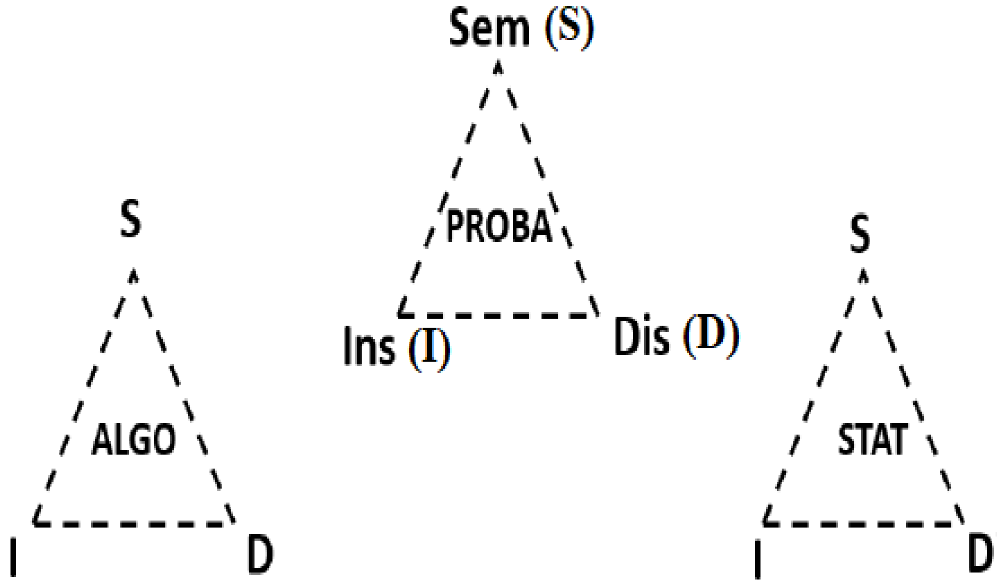
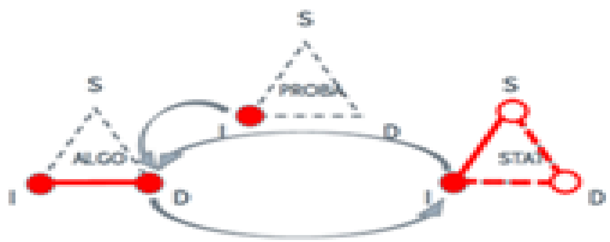



**Fig. 2 :** Vue de dessus des « projections » dans les ETM et ETA

```
DÉBUT
  TEST ← 0
   SUCCES ← 0
   TANTQUE TEST ≤ 400 FAIRE
      TEST ← TEST + 1
      SI ABS(ALEA() – ALEA()) > 0,5 ALORS
         SUCCES ← SUCCES + 1
      FINSI
   FINTANTQUE
   FREQUENCE ← SUCCES/400
   AFFICHER FREQUENCE
FIN
```

Algorithme de niveau A1

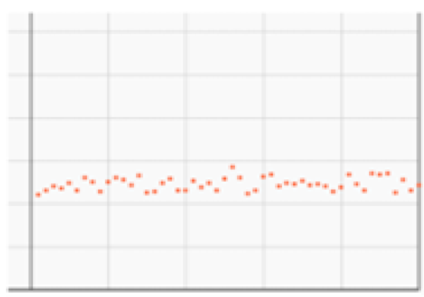


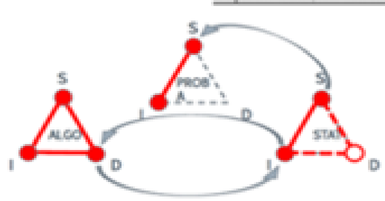

Analyse des 3 niveaux de l'algorithme A1

**Fig. 3a :** Algorithme de niveau A1

```
DÉBUT
   K ← 0
   TANTQUE k ≤ 50 FAIRE
      EXÉCUTER PROGRAMME NIVEAU A1
      k ← k + 1
      TRACER_POINT (k ; FREQUENCE)
   FINTANTQUE
FIN
```

Algorithme de niveau A2






Analyse des 3 niveaux de l'algorithme A2

**Fig. 3b :** Algorithme de niveau A2 après implémentation dans un environnement numérique

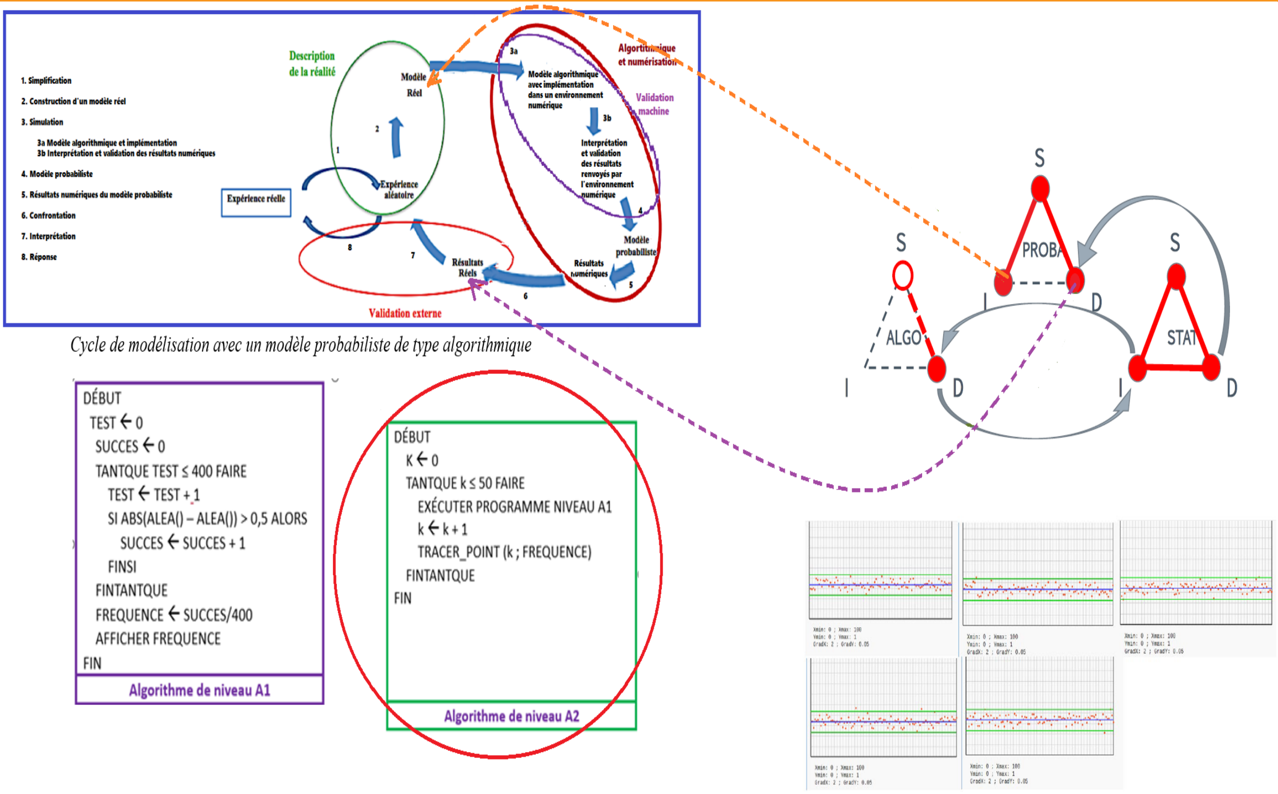



**Figure 4:** Lien entre les projections ETM/ETA et le cycle de modélisation avec modèle probabiliste de type algorithmique

## Références


Douady R. (1986) Jeux de cadres et dialectique outil-objet. *Recherches En Didactique Des Mathématiques, 7*(2), 5–31.

Duval R. (1993) Registres de représentation sémiotique et fonctionnement cognitif de la pensée. *Annales de Didactique et de Sciences Cognitives, 5*, 37–65.

Kuzniak A. (2011) L'Espace de Travail Mathématique et ses genèses. *Annales de Didactique et de Sciences Cognitives, 16*, 9–24.

Kuzniak A., & Richard P. (2014) Spaces for mathematical work: Viewpoints and perspectives. *Relime, 17*(4.1), 17–26.

Lagrange J.-B. & Laval D. (2019) Connected Working Spaces: the case of computer programming in mathematics education. *Eleventh Congress of the European Society for Research in Mathematics Education*. Utrecht, Netherlands. https://hal.archives-ouvertes.fr/hal-02418181 (Consulté le 29/05/2023)

Laval D. (2018) L'algorithmique au lycée entre développement de savoirs spécifiques et usage dans différents domaines mathématiques. Thèse de Doctorat. Université Sorbonne Paris Cité.

Laval D. (2021) Les Espaces de Travail Mathématique et Algorithmique connectés. Étude d'une activité algorithmique comme objet d'apprentissage de savoirs spécifiques et d'usage : Cas d'une ingénierie « dichotomie continue ». In A. Chesnais, H. Sabra (Eds.), *Actes du séminaire national de didactique des mathématiques 2020* (p.88-114). IREM de Paris - Université de Paris. https://hal.archives-ouvertes.fr/hal-03545511/ (Consulté le 29/05/2023)

Nechache A. (2016) La validation dans l'enseignement des probabilités au niveau secondaire. Thèse de doctorat. Université Denis Diderot Paris 7.

Parzysz B. (2009) Des expériences au modèle, via la simulation. In *Repères-IREM*, (74) : 91–103.

Rabardel P. (1995). Les hommes et les technologies ; approche cognitive des instruments contemporains. Armand Colin, pp.239. https://hal.archives-ouvertes.fr/hal-01017462 (Consulté le 29/05/2023)

Trunkenwald J. (2018) L'Espace de Travail Mathématique : Une respiration pour analyser la fluctuation d'une simulation d'échantillonnage. *Mathematical Working Space, Proceedings Sixth MWS Symposium*. Valparaiso, Chili. New submission ETM 6.

Trunkenwald J. & Laval D. (2019) Algorithms as a discovery process in frequentist approach to prediction interval. *Eleventh Congress of the European Society for Research in Mathematics Education (CERME11)*, Utrecht, Netherlands. https://hal.archives-ouvertes.fr/hal-02412851 (Consulté le 29/05/2023)

Eduscol – Mathématiques Lycée. (2009) Ressources pour la classe de seconde, Probabilités et statistiques. https://maths.ac-noumea.nc/IMG/pdf/Probabilite_s_et_statistiques_1.pdf (Consulté le 29/05/2023)