

Sujet TRAAM : « La famille Agache triche-t-elle pour la galette des rois ? »

Public : Terminales Générales en ECS

Plan de travail :

I. Phase exploratoire :

DHC sur le modèle de ceux donnés en 1G ayant pour question : « La famille Agache triche-t-elle pour la galette des rois ? »

- a. Diffusion de deux vidéos réalisées par les 1G suivantes :

[DHC Pamplermousse](#)

[DHC Distanciation](#)

- b. Discussion sur l'aspect modélisation et distribution des éléments de correction et d'évaluation suivants :

[Eléments de correction](#)

- c. Le DHC : Pour ce DHC, vous devez réaliser une courte vidéo (impérativement moins de 5 min) pour présenter ce sujet, issu de la vie « réelle », ainsi que la façon de le modéliser en mathématiques. L'objectif n'est pas de résoudre la question mais plutôt d'envisager comment on pourrait répondre à une telle question (quel(s) modèle(s), quelle(s) hypothèse(s) simplificatrice(s), de quelle(s) information(s) doit on disposer...). Notez également qu'il s'agit essentiellement d'un exercice d'oral pour lequel je ne vous impose rien du point de vue de la réalisation.

Ressource : Fichier Excel/Word pour évaluation formative ([Non fournis](#))

Commentaires prof : L'idée ici est à la fois de commencer par un modèle simple (hasard ou non) mais aussi d'espérer voir émerger des stratégies de la part des élèves (les intervalles de fluctuation et de confiance ainsi que les probabilités conditionnelles ont déjà été revu à l'occasion du cours sur la biodiversité)

A l'issu de ces devoirs, on convient des hypothèses de travail, des notations. On évoque également les biais de cette première modélisation probabiliste.

Modèle retenu pour la suite : expérience aléatoire.

On va estimer que l'objectif d'un tricheur est de favoriser le plus jeune.

On note T l'événement la famille triche, J l'événement le plus jeune a la fève.

- Si la famille ne triche pas : La probabilité de J est q
- Si la famille triche : La probabilité de J est r

Hypothèses simplificatrices : le comportement de la famille est considéré comme stable (tricheur ou non). Sa variation éventuelle est en fait prise en compte dans r .

II. Approche probabiliste :

DHC très court suivi d'une correction collective.

On estime que $q=1/4$. On a les résultats suivants pour deux familles.

Ressource : Fichier Excel ([Simulation](#))

Pensez vous que ces familles ont triché ? On justifiera sa position par des arguments de nature mathématique.

Proposez, sous la forme d'une courte vidéo ou d'un simple fichier audio, une façon de déterminer si la famille Agache triche. On prendra soin de préciser le modèle mathématique utilisé ainsi que ses biais éventuels.

Commentaires prof : intervalle de fluctuation, notion de grand nombre. Ici le travail doit être très court car la notion a été revue et figure au programme de seconde. L'idée est de donner lors de la correction des éléments facilitant l'approche suivante.

III. Approche statistique :

Travail en classe sous la forme de 3 sujets différents avec une phase de mutualisation

Sujet 1 (différenciation : lvl 1)

Nous vous avons donné deux réponses au choix : « Oui bien sûr » et « Non, c'est sacrilège ». Surprise : sur les quelque 8 000 lecteurs à voter, vous êtes 53 % à penser que c'est sacrilège de tricher ! Nous qui étions sûrs que dans toutes les familles du monde, on donnait la fève au petit dernier !

La Voix du Nord 02/01/2020

- Quel modèle mathématique est utilisé ici ? Sur quel principe repose-t-il ?
- Quelle information peut-on réellement établir à partir d'un tel sondage ?
- Quels sont les biais éventuels ?

Sujet 2 (différenciation : lvl 2)

Les français seraient-ils mauvais joueurs ? Lors de la prochaine **épiphanie**, méfiez-vous lors de la distribution des parts de la galette des rois : nous serions 79% à tricher pour la fève !

**Méthodologie : enquête réalisée entre le 14 et 24 décembre 2020 auprès de 11 316 personnes représentatives de la population nationale française selon la méthode des quotas. Sondage effectué en ligne, sur le panel propriétaire BuzzPress France regroupant 24 802 personnes. Toutes les informations mises en avant par les personnes interrogées sont déclaratives.*

- Quelle méthode a été utilisée pour établir cette information ?
- Quels problèmes pose cet énoncé d'un point de vue mathématique ?
- Proposez une autre exploitation mathématique des données de cette enquête.
- Quels sont les biais éventuels qui subsistent ?

Sujet 3 (différenciation : lvl 3)

Ce sujet est composé des deux extraits de presse précédent ainsi que du suivant :

Même si 68% des Français reconnaissent volontiers tricher pour que les enfants la trouvent, la miniature reste un incontournable de la fête. Décryptage de ce marché florissant de la galette des Rois.

- Pour chaque article, déterminez les difficultés qu'il pose d'un point de vue mathématique.
- Proposez un modèle mathématique permettant d'établir les trois indicateurs statistiques donnés dans ces articles.
- Les deux premiers sujets donnent-ils des informations cohérentes ?
- Pour quelle taille d'échantillon l'article 3 est-il cohérent avec l'article 2 ?

Mutualisation puis question commune : Proposer une nouvelle façon de déterminer si la famille Agache triche. On prendra soin de préciser le modèle mathématique utilisé ainsi que ses biais éventuels.

Commentaires prof : Intervalle de confiance, notion de grand nombre. Revenir sur la différence entre ces deux expériences mais sur l'équivalence mathématique. Première comparaison de modèles. Importance de l'esprit critique dans les programmes d'ECS.

IV. Inférence Bayésienne :

Devant la difficulté à obtenir une véritable étude statistique, on va supposer que la probabilité que la famille Agache triche est de 0,5 (modèle équiprobable). On va également considérer, pour simplifier que $q = 1/4$ et que $r=3/4$

1. On sait à présent que c'est bien le plus jeune qui a obtenu la fève lors de la première galette de l'année 2021. En ayant cette information, expliquez pourquoi on peut considérer que la probabilité que la famille Agache triche est à présent de 0,75 (En réalité, c'est la probabilité sachant J).

Différenciation :

- 1G : arbre de probabilité conditionnelle
 - Non 1G : tableau à double entrée
2. Lors d'une deuxième galette, justifiez pourquoi on préfère considérer que $p(T)=0,75$ plutôt que $p(T)=0,5$.
 3. Vérifiez que si pour la deuxième galette ce n'est pas le plus jeune qui a la fève alors on considère à nouveau que $p(T)=0,5$.

DHC final : pour des « familles données » comparer les différents modèles et répondre à la question initiale sous la forme d'une vidéo capsule de 5min max.

Ressource : Fichier EduPython ([Inférence bayésienne](#))

V. Réinvestissements :

- a. Les tests médicaux (programme ECS) : sur la prévalence et comment l'établir

FIG 1	Rendement diagnostique du frottis nasopharyngé-PCR Covid-19
--------------	--

Rendement diagnostique du frottis nasopharyngé-PCR pour la maladie Covid-19 selon la prévalence supposée de la maladie au sein de la population dont est issu le patient (probabilité prétest)

VPP: valeur prédictive positive = probabilité d'avoir la maladie sachant que le test est positif; VPN: valeur prédictive négative = probabilité de ne pas avoir la maladie sachant que le test est négatif.

Prévalence 10%

	Maladie + (100)	Maladie - (900)	Spécificité	0,99
Test +	56 - 83	9	Sensibilité	0,56-0,83
Test -	17 - 44	891	VPP	0,86 – 0,90
			VPN	0,95 – 0,98

Prévalence 20%

	Maladie + (100)	Maladie - (400)	Spécificité	0,99
Test +	56 - 83	4	Sensibilité	0,56-0,83
Test -	17 - 44	396	VPP	0,93 – 0,95
			VPN	0,90 – 0,96

Prévalence 30%

	Maladie + (100)	Maladie - (233)	Spécificité	0,99
Test +	56 - 83	2	Sensibilité	0,56-0,83
Test -	17 - 44	231	VPP	0,97 – 0,98
			VPN	0,84 - 0,93

Prévalence 50%

	Maladie + (100)	Maladie - (100)	Spécificité	0,99
Test +	56 - 83	1	Sensibilité	0,56-0,83
Test -	17 - 44	99	VPP	0,98 – 0,99
			VPN	0,69 – 0,85

Résumé :

Probabilité Pré-test	VPN
10%	0,95 - 0,98
20%	0,90 - 0,96
30%	0,84 - 0,93
50%	0,69 - 0,85

Rev Med Suisse 2020; 16: 699-701

Ressources : Fichier EduPython ([Programme sur la succession de test](#))

- b. L'I.A : exemple conseillé des SPAM qui permet de lier les approches fréquentielles (pour déterminer un SPAM WORD) et l'inférence Bayésienne (pour une adaptation locale)

Ressources professeur : Deux vidéos capsules ([ici](#) et [là](#) au format pdf) pour comprendre les articulations didactiques rédigées avec pour public tous les enseignants d'ECS.