

https://www.vauche.com/ https://www.vauche-occasions.com/

Master 2 : Ingénierie Statistique Numérique - Data Sciences

Stage: Entreprise Vauché

Estimation qualité tri des déchets - Société Vauché

Réalisé par :

Youssouf Roki-Dine

Sous la direction de Céline Duval Stéphane Vauché

Année universitaire : 2023-2024

Remerciements

Table des matières

I	Histor	orique des Établissements						
II	Introd	luction	4					
	II.1	Contexte et Expertise de Vauché						
	II.2	Objectif du Cahier des Charges	5					
	II.3	Problématique : Maximiser la Pureté tout en Valorisant les Déchets	5					
III	Présentation du problème d'échantillonnage							
	III.1	Plan de sondage aléatoire	7					
	III.2	Classification des Plans de Sondage aléatoire	7					
	III.3	Taille d'échantillon	7					
IV	Exploration des Données du centre de tri d'Angers							
	IV.1	Plan d'échantillonnage par grappe (pour vauché)	9					
	IV.2	Analyses des échantillons de flux sortants	10					
	IV.3	Constitution des échantillons en entrant et répartition	13					
V	Amélioration des Techniques d'Échantillonnage pour une Évaluation Fiable							
	du Tri des Déchets							
VI	Simulation							
	VI.1	Résultats Numériques	18					
	VI.2	Analyse des Taux de Confiance par Catégorie Maj. et Min	30					
VII	Annexe							
	VII.1	Modélisation jouet du problème	36					
	VII.2	Déclaration des paramètres pour le cas jouet	37					
	VII.3	Estimateur des paramètres	38					
	VII.4	Loi de l'estimateur du paramètre	39					
	VII.5	Intervalle de Confiance	43					
	VII.6	Les graphes pour Décembre	44					
	VII.7	Les graphes pour Janvier 2002	50					
	VII.8	Les graphes pour Février	57					

I Historique des Établissements

La Maison Vauché est une entreprise de chaudronnerie qui a su relever, au fil des décennies, de nombreux défis, survivre à trois occupations désastreuses - et pas n'importe où, à Sedan, « trouée » historique, traditionnel couloir d'invasion -, traverser de multiples récessions au cœur d'un vieux bassin industriel, s'adapter aux nouvelles demandes du marché, changer, se métamorphoser, en un mot : innover. De 1860 à 2010, le groupe Vauché est passé de l'ère brassicole à celle de l'Environnement avec un grand « E ». Plusieurs générations successives d'une même famille, originaire des rives mosanes et des contreforts de l'Ardenne, ont travaillé d'arrache-pied pour faire perdurer une grande marque de qualité, comme le fait remarquer à juste titre, l'illustre préfacier Frédérick Tristan, lauréat du Prix Goncourt. A la suite d'interviews, de prospections, de nombreuses recherches dans des fonds privés et publics d'archives, l'historien Gérald Dardart, ancien élève de Lettres supérieures au Lycée Henri-IV et docteur en histoire moderne de l'université de Paris-Sorbonne, auteur de dizaines d'ouvrages et articles, vous propose ici d'analyser, épaulé par une présentation rigoureuse de documents, les différentes périodes charnières, l'évolution du Groupe Vauché et son art d'adaptation face aux crises et aux bouleversements du marché. Gérald Dardart s'est toujours appliqué à resituer l'entreprise dans un contexte, à la fois, politique, juridique, industriel. Disponible sur cette adresse [4]

II Introduction

II.1 Contexte et Expertise de Vauché

Vauché est un acteur majeur dans le domaine du tri des déchets. L'entreprise se positionne en tant que concepteur et constructeur d'usines de déchets, offrant ses services aussi bien au secteur privé qu'au secteur public. Leur expertise se concentre sur le tri de déchets, reconnaissant que chaque maillon de tri représente un ensemble complexe de trois étapes essentielles : la collecte, le tri, et le recyclage :

1. Collecte des Déchets:

- Vauché récupère les déchets provenant des collectivités locales, tels que les poubelles de recyclage.
- Ces déchets sont acheminés vers le centre de tri, où ils sont préparés pour le processus de tri.

2. Tri Automatique:

• Dans le centre de tri, des systèmes automatisés trient les matériaux en fonction de leurs propriétés physiques (taille, densité, etc.).

• Le tri automatique permet de séparer les cartons, le verre, les plastiques et autres matériaux recyclables.

3. Cahier des Charges:

- En sortie, l'usine doit assurer un certain cahier des charges. Cela signifie que les matériaux triés doivent répondre à des critères spécifiques de pureté.
- Par exemple, si l'usine traite 100 tonnes de déchets, elle doit garantir que la tonne de carton extraite est de haute qualité, sans contamination excessive.

II.2 Objectif du Cahier des Charges

Dans le cadre de la construction d'une nouvelle usine de tri, des critères rigoureux doivent être respectés. Ces critères garantissent que l'usine fonctionne de manière optimale, tout en minimisant son impact environnemental. Voici un exemple pour illustrer ces exigences :

Exemple de Critère : Capacité de Traitement Quotidienne

- L'usine doit être capable de traiter entre 8 et 15 tonnes de déchets par heure, soit entre 112 et 210 tonnes par jour, ce qui représente presque une moyenne de 12 tonnes par heure.
- Cette capacité est essentielle pour répondre aux besoins croissants de la région en matière de gestion des déchets.

II.3 Problématique : Maximiser la Pureté tout en Valorisant les Déchets

La complexité réside dans la balance entre pureté et valorisation maximale des déchets. Voici les défis :

• Population Pléthorique :

- Imaginez qu'on interroge 100 personnes pour comprendre le comportement de 100000 personnes. De même, Vauché doit échantillonner les déchets pour garantir la qualité de l'ensemble.
- Comment peut-on extraire 1000 kilogrammes de carton de haute pureté à partir de 2000.000 tonnes de déchets hétérogènes?

• Contraintes de Pureté et de Temps :

- L'usine doit respecter des contraintes de pureté (par exemple, 95% de carton pur) tout en minimisant le temps de traitement.
- Valoriser une seule tonne de carton de haute qualité parmi 100 tonnes de déchets est le défi.

III Présentation du problème d'échantillonnage

Pour évaluer la qualité du tri des déchets, il est nécessaire de comprendre les caractéristiques statistiques du tri effectué sur une certaine période ou quantité de déchets. Cette évaluation peut être réalisée par deux approches principales :

- Le **recensement**, qui serait l'approche descriptive, impliquerait de passer en revue et d'évaluer chaque article de déchet trié dans la période étudiée. Cela garantirait une précision maximale, mais s'avère souvent impraticable en raison du temps ou des coûts élevés associés.
- L'échantillonnage, en revanche, est une approche inférentielle utilisée lorsque le recensement n'est pas réalisable. Il consiste à sélectionner et analyser une partie représentative des déchets triés, nommée échantillon. Si l'échantillon est choisi correctement, les caractéristiques statistiques de cet échantillon offriront une estimation fiable de la qualité du tri de l'ensemble du volume de déchets.

Prenons l'exemple de la volonté de connaître la qualité du tri des déchets pour l'ensemble des déchets traités en décembre :

- Recensement : cela signifierait d'évaluer chaque article de déchet trié en décembre, ce qui est extrêmement minutieux et probablement trop coûteux et long à réaliser.
- Échantillonnage : nous sélectionnerions une partie représentative des déchets de décembre, et nous analyserions la qualité de cette portion en espérant qu'elle reflète fidèlement la qualité globale du tri pour tout le mois.

Cette méthode soulève des questions importantes pour l'estimation de la qualité du tri :

- Comment choisir l'échantillon représentatif des déchets?
- Quelle quantité de déchets devrait constituer l'échantillon?
- Comment assurer que l'échantillon est représentatif de la diversité des déchets?

La réponse à ces questions implique une compréhension approfondie de la composition des déchets et des processus de tri, ainsi que l'application de techniques statistiques pour garantir que l'échantillon sélectionné fournisse une estimation valide de la qualité du tri. Il est également crucial d'évaluer l'erreur potentielle liée à l'échantillonnage et de déterminer si cette estimation est satisfaisante pour les besoins de la société. Par exemple, nous pourrions viser à ce que l'évaluation de la qualité du tri des déchets pour le mois de décembre soit représentative avec un intervalle de confiance précis, comme une marge d'erreur acceptable fixée par la société.

III.1 Plan de sondage aléatoire

Le sondage aléatoire simple est fondé sur le principe d'équité et d'impartialité où chaque échantillon de taille n dans une population totale de N a une chance égale d'être sélectionné. c'est ce que proposent par exemple la théorie des sondages proposer par le professeur (Laurent Rouvière) de l' Université Rennes [1].

III.2 Classification des Plans de Sondage aléatoire

Un plan de sondage est avec remise, losque chaque individu sélectionné est remis dans la population et peut donc être choisi à nouveau. L'ordre de sélection est également pris en compte. Le nombre total de combinaisons possibles est de N^n , puisque chaque élément peut être choisi n fois de manière indépendante.

À l'opposé, un plan de sondage sans remise ne permet pas qu'un individu soit sélectionné plus d'une fois, supprimant ainsi les doublons dans les échantillons. Ce type de sondage est représenté par les combinaisons et leur nombre est donné par la formule du coefficient binomial :

$$\binom{N}{n} = \frac{N!}{n!(N-n)!}$$

Ce concept est crucial en statistique puisqu'il affecte directement la variance des estimateurs et les intervalles de confiance.

Chaque échantillon dans un plan simple a une probabilité $1/\binom{N}{n}$ d'être tiré. Cela garantit que tous les échantillons sont équiprobables, une condition nécessaire pour l'inférence statistique valide.

III.3 Taille d'échantillon

Il existe différents facteurs à considérer dans la détermination de la taille de l'échantillon souvant la taille d'échantillon est fixe. Cependant, la question se pose souvent pour nous les statisticiens : "À partir de combien d'éléments un échantillon devient-il fiable?" on se réfère à la précision des estimations que nous pouvons obtenir à partir de cet échantillon. Dans notre contexte, cela implique de convenir d'un écart maximal toléré pour l'intervalle de confiance. Autrement dit, nous cherchons la taille d'échantillon minimum n_0 telle que l'intervalle de confiance ne dépasse pas une certaine largeur prédéterminée. Plus précisément, nous souhaitons déterminer la demi-largeur h_0 de l'intervalle de confiance telle que la taille d'échantillon n_0 soit suffisante pour garantir que la demi-largeur de l'intervalle de confiance soit égale à h_0 . [1]

IV Exploration des Données du centre de tri d'Angers

Pour répondre à la problématique de l'estimation de la qualité du tri des déchets, nous avons entrepris une série d'analyses de données basées sur des essais réalisés au centre de tri d'Angers. L'objectif de ces analyses est de comprendre en profondeur les caractéristiques du flux de déchets et l'efficacité des processus de tri mis en œuvre.

Répartition des tonnages entrants par mois

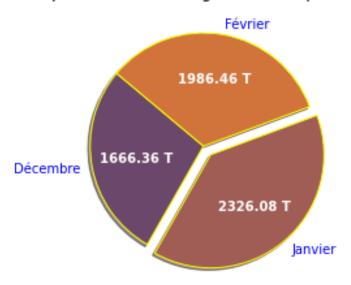


FIGURE 1

Ce graphique illustre la répartition des tonnages de déchets à traiter par le centre de tri d'Angers au cours de trois mois consécutifs. Chaque segment représente un mois différent, avec une masse totale de déchets exprimées en tonnes.

En décembre, la quantité traitée était de 1666,360 tonnes. Ce mois semble être celui avec le volume le plus bas sur la période observée, ce qui pourrait s'expliquer par les interruptions liées aux fêtes de fin d'année ou par des variations saisonnières dans la production de déchets.

Janvier a vu une augmentation significative, avec un total de 2326.080 tonnes de déchets traités, comme le montre dans ce graphique. Cette augmentation pourrait être due à une accumulation de déchets post-fêtes ou à la reprise des activités économiques après les congés de décembre.

Enfin, le mois de **février**, correspondant au segment de couleur orange, a enregistré le tonnage le plus élevé avec **1986.46 tonnes**. Bien que février soit un mois plus court,

la quantité de déchets traités est restée relativement élevée, suggérant une production constante de déchets ou une efficacité accrue dans les opérations de traitement.

Par exemple voici un tableau qui présente une vue d'ensemble de données des tonnages de déchets traités par l'usine durant le mois de décembre 2021.

		т	ONNAG	ES EN	TRANT	S + REPAR	TITION TH	IEORIQUE		
déc-21	ALM	BV	L&S	SMIPE	LOIRE B	SAUMUR VL	SUD SAUM	TOTAL		
dec-21	MULTI	MULTI	MULTI	ЕМВ	ЕМВ	ЕМВ	ЕМВ	MULTI	ЕМВ	TOTAL
STOCK A TRIER	769,700	205,780	153,560	45,860	116,040	257,260	118,160	1129,040	537,320	1666,360
		OBJECTIFS			ОВЈІ	ECTIFS		OBJE	CTIFS	OBJECTIF
EMR	201,652	40,409	41,067	12,791	33,502	80,679	42,399	283,128	169,372	452,499
JRM (I.II)	221,254	77,947	44,242	2,885	4,624	27,075	9,059	343,443	43,643	387,086
Autres refus (incluant fines < 50mm)	137,478	17,304	22,666	8,739	22,663	54,007	15,026	177,448	100,435	277,883
PET Clair Bouteilles / flacons (Code	32,147	8,748	7,406	4,764	9,492	14,731	7,602	48,301	36,589	84,889
ACIER	25,075	11,225	6,865	4,427	8,818	17,014	10,253	43,165	40,512	83,677
Pots et Barquettes PE/PP (Codes 2 -	28,968	8,942	5,863	1,700	7.450	15,455	8,795	43,774	33,401	77,175
Gros de magasin (= fibreux < 120 mi	33,436	8,953	5,284	0,776	1,905	2,862	0,690	47,673	6,232	53,905
PEHD Bouteilles et Flacons (Code 2)	12,185	5,502	3,520	2,216	3,960	7,433	5,367	21,207	18,977	40,184
ELA	14,378	7,641	3,069	1,689	5,787	6,438	3,165	25,089	17,079	42,168
Films PE (Code 4)	12,922	5,214	3,286	1,434	3,379	9,418	3,986	21,422	18,217	39,639
PET Couleur Bouteilles / flacons (Flu	8,863	4,085	2,032	1,329	2,651	6,717	3,099	14,980	13,796	28,776
Films refusés (PP, composites) (Co	12,487	0,478	1,252	1,094	3,607	1,283	1,540	14,217	7,524	21,741
Verre	8,298	2,987	2,198	0,406	3,466	3,712	1,945	13,484	9,529	23,012
ALU	6,550	1,750	1,228	0,508	1,056	2,720	1,115	9,528	5,399	14,927
PS (Flux développement) (Code 6)	2,838	1,351	1,317	0,460	1,388	2,715	1,597	5,506	6,160	11,666
Barquettes monocouche PET clair et	6,183	1,789	1,054	0,172	1,288	2,428	1,663	9,025	5,552	14,577
Pots Barquettes refusés (Barquettes	2,486	0,726	0,581	0,131	0,358	0,999	0,217	3,793	1,705	5,497
ACIER IMBRIQUES	1,016	0,173	0,351	0,048	0,276	1,147	0,283	1,540	1,754	3,295
PSE (Flux développement) (Code 6)	0,701	0,330	0,136	0,226	0,302	0,142	0,066	1,167	0,736	1,903
PSE barquettes noires	0,782	0,225	0,142	0,064	0,069	0,286	0,293	1,150	0,711	1,861

FIGURE 2 – Répartition des déchets Déc. 2021

Les données du tableau représentent l'intégralité des déchets à traiter par la société au mois de décembre, avec un tonnage global de 1 666 360 kg. Afin d'effectuer une évaluation approfondie et efficace de la qualité du tri, un échantillon a été sélectionné de cette population totale.

IV.1 Plan d'échantillonnage par grappe (pour vauché)

On constate que les objets sont naturellement regroupés (par exemple, par lot de production), c'est dans cette perspective que Vauché sélectionne aléatoirement quelques grappes et inclure tous les objets de ces grappes dans l'échantillon. Voici un exemple qui illustre un grappe de flux entrant qui constitue l'échantillon (voir Figure 3)

Sous-catégorie 🔻	Masse brute (k ▼	Tare (kg)	Masse nette (k ▼	% éch. ▼
Flux Développement	16,30	9,10	7,2	14,66%
PEHD/PP	17,20	8,80	8,4	17,11%
ELA	1,20	0,55	0,65	1,32%
Aluminiums	0,85	0,55	0,3	0,61%
Aluminiums souples	0,55	0,50	0,05	0,10%
Papiers Cartons Mélés (= gros de magasin) : fibreux > 120x50	6,65	6,05	0,6	1,22%
Papiers (1.11)	28,25	16,80	11,45	23,32%
Acier	1,65	0,50	1,15	2,34%
EMR	25,80	16,85	8,95	18,23%
Film PEBD	9,10	8,30	0,8	1,63%
Refus	23,45	17,10	6,35	12,93%

FIGURE 3 – Une grappe de Flux Entrants

En décembre, la quantité traitée s'élevait à 1666,360 Tonnes. Nous avons constaté que trois échantillons de flux entrants ont été sélectionnés aléatoirement, avec une masse nette de 161,47 kg par jour. Cela nous amène à conclure qu'un échantillon représentatif de 3542 kg par mois devrait être sélectionné parmi l'ensemble des déchets stockés à trier pour décembre 2021.

En janvier, une augmentation significative a été observée, avec un total de 2329.080 Tonnes. traitées. Trois échantillons de flux entrants ont été sélectionnés aléatoirement, avec une masse nette de 151,65 kg par jour. Cela nous amène à conclure qu'un échantillon représentatif de 3336,3 kg par mois devrait être sélectionné parmi l'ensemble des déchets stockés à trier pour Janvier 2022.

On observe la même analogie sur le mois de Février.

IV.2 Analyses des échantillons de flux sortants

Décembdre:

Suite à l'étude réalisée avec un échantillon représentatif des déchets traités en décembre, nous observons des résultats significatifs en termes de tri correct en sortie de cabine.

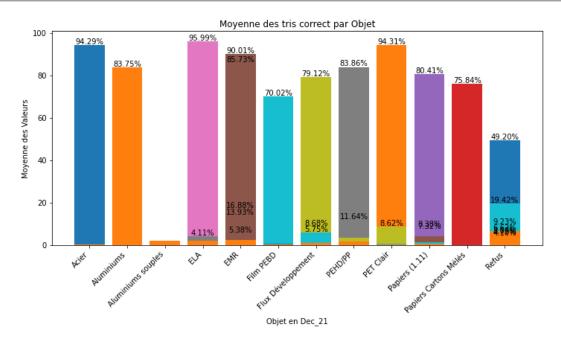


Figure 4

Le graphique montre que l'Acier, ELA et PET clair ont les taux de tri correct les plus élevé respectivement avec 94.29%, 95.99%, et 94.31%, indiquant un tri extrêmement précis pour ces objets. EMR présente également un taux élevé de tri correct à 90.01%.

Pour les objets comme le Film PEHD, Flux Développement et Papiers Cartons Mêlés, leurs taux de tri correct sont respectivement de 70.02%, 79.12% et 75.84%, suggérant que bien qu'ils soient majoritairement bien triés pourtant, il reste une marge très importante pour améliorer le tri de ces matières. Le Films PEBD refusés, qui atteint 19.42%, et en termes de masse moyenne de l'échantillon, il représente la masse la plus faible, cela pourrait refléter la présence des éléments non-conformes, tels que les téléphones, batteries, DEEE, morceaux de bois de grande taille, conserves pleines et aérosols, qui nécessitent des processus de traitement distincts et ne doivent pas être intégrés dans les flux de recyclage standards. Ces objets légers et souvent fins peuvent être plus difficiles à séparer efficacement lors du processus de tri, entraînant un taux de rejet supérieur. Le PEHD/PP et Papiers montrent des résultats assez positifs avec des taux de 83.86% et 80.41%, ce qui est encourageant pour ces flux de recyclables.

On peut observer ici que dans la catégorie des Refus, avec seulement 49.20% vrai refus et 50.8% d'autre objets valorisables, Cela indique que presque la moitié de ces déchets supposer être de refus sont valorisables, soulignant un domaine potentiel d'intervention pour augmenter l'efficacité du tri et réduire la contamination.

Ces résultats démontrent l'importance de l'échantillonnage dans la mesure des performances de tri et mettent en lumière les succès ainsi que les défis à relever pour optimiser le processus de tri des déchets.

Janvier:

Pour le mois de janvier, l'analyse de l'échantillon sélectionné montre une moyenne des taux de tri corrects variable selon les différents objets.

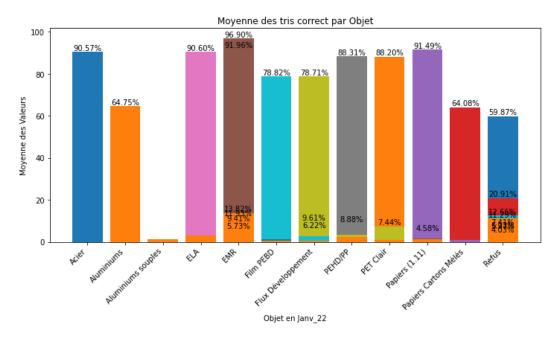


FIGURE 5

L'acier reste bien trié avec un taux de 90.57%, tandis que les aluminiums ont connu une baisse notable avec un taux de 64.75% de tri correct, ce qui est inférieur à celui observé en décembre. Cette diminution pourrait indiquer des défis spécifiques rencontrés avec le tri de ces matériaux pendant ce mois.

Les taux de réussite pour l'ELA et l'EMR sont respectivement de 90.60% et (91.96% et 96.90%), montrant une amélioration par rapport à décembre, ce qui suggère une optimisation dans le traitement de ces flux. Par contre, le film PEHD affiche un taux de 78.82%, et le Flux Développement est à 78.71%, tous deux reflétant une marge d'amélioration dans le tri de ces plastiques.

Concernant le PEHD/PP et le PET clair, les taux sont respectivement de 88.31% et 88.20%, indiquant une constance dans le tri correct de ces matériaux par rapport au mois précédent. Pour Papiers (1.11) un taux de 91.49% illustre une très bonne qualité de

tri, tandis que Papiers Cartons Melés, sont légèrement moins bien triés avec un taux de 64.08%.

Enfin, le taux de refus, à 59.87%, reste significatif, suggérant que presque 40% d'autres objets valorisants de cette catégorie ont été triés.

Février:

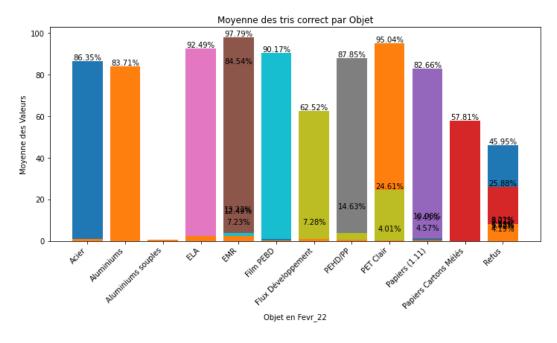


FIGURE 6

IV.3 Constitution des échantillons en entrant et répartition

Suite aux résultats observés, nous sommes en mesure de reconstituer des échantillons d'entrants pour chaque mois à partir des analyses de flux de sortie présentées dans la section précédente. Dans le cadre de ces analyses, un échantillon représentatif de 1147 kg a été sélectionné aléatoirement parmi la masse totale des déchets traités de 1666.360 tonnes, et cette échantillon est répartie de façons inégale comme le montre le Figure 7.

Répartition de la masse d'échantillon (Dec_21)

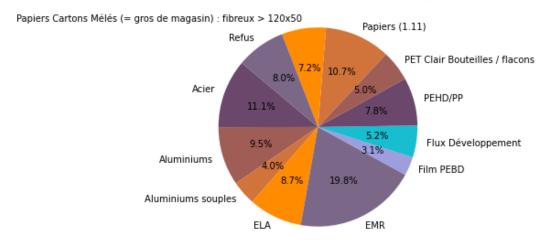


FIGURE 7

Les sections du graphique illustrent la diversité de type d'objet dans l'échantillon, avec des pourcentages correspondant à leur masse respective. Par exemple, le segment le plus large, représente les EMR, qui constituent 19.8% de l'échantillon. Cette catégorie inclut des matériaux tels que les gros de magasin, signifiant des déchets en fibres recyclables d'une dimension standardisée.

D'autres matériaux, comme le "Papiers(1.11)" et l' "Acier", sont également significatifs dans l'échantillon, avec respectivement 10.7% et 11.1% de la masse. Ces pourcentages révèlent non seulement la composition des déchets triés mais aussi la potentialité de recyclage de chaque matériau.

L'utilisation d'un échantillon est une méthode inférentielle qui permet d'estimer les caractéristiques de l'ensemble de la masse des déchets sans avoir à effectuer un recensement, souvent coûteux et peu pratique. L'échantillon de 1242.51 kg constitue environ 0.07% de la masse totale du mois, ce qui est une taille d'échantillon suffisamment représentative pour donner une estimation fiable de la composition globale des déchets, à condition que l'échantillonnage soit réalisé de manière aléatoire et systématique.

Cette analyse est cruciale pour évaluer l'efficacité du tri et pour identifier les opportunités d'optimiser le processus de recyclage.

Dans cette paragraphe, constatons que pour évaluer la qualité du tri des déchets, un échantillon représentatif de 1150.84 kg a été examiné, par rapport à la masse totale

traitée de 2 326 .080 Tonnes.



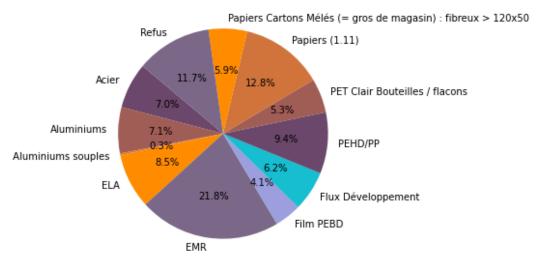


FIGURE 8

Le segment qui représente les papiers(1.11) et constitue 12.8% de l'échantillon, indiquant une part significative des matériaux facilement recyclables. Le "PET Clair Bouteilles/flacons" représente 5.3%, ce qui est important pour les processus de recyclage du plastique.

Les "Aluminiums" et les "Aluminiums souples" forment ensemble près de 7.4% de l'échantillon, ce qui est non négligeable et souligne l'importance de séparer correctement ces objets qui sont hautement recyclables. Notamment, les "Aluminiums souples" représentent 0.3% de l'échantillon, une quantité petite mais significative en raison de la valeur du recyclage de l'aluminium.

Le "Flux Développement" et le "Film PEBD" ont des parts respectives de 6.2% et 4.1%, suggérant une présence notable de plastiques divers au sein du flux de déchets, avec des implications pour les opérations de tri spécifiques à ces objets.

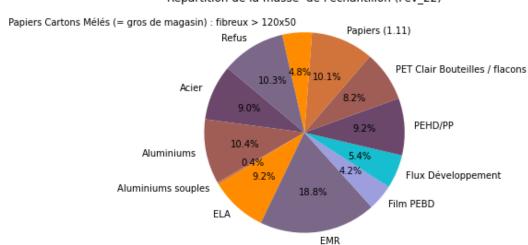
Le "Refus", qui comprend les matériaux ne pouvant pas être recyclés ou contaminés, occupe 11.7% de l'échantillon. Cette proportion est critique, car elle indique la quantité de déchets qui ne sont pas valorisable dans le processus de tri actuel et pourrait orienter les efforts vers la réduction de ce taux.

Les "EMR", le plus volumineux, représente 21.8% de l'échantillon et suggère une quantité significative de déchets potentiellement valorisables, nécessitant une attention particulière pour optimiser leur récupération.

En somme, l'échantillon de 1150.84 kg, qui représente approximativement 0.05% de la masse totale des déchets de janvier, fournit une fenêtre précise sur la composition des déchets triés pendant ce mois. Ces données servent de point de départ pour analyser l'efficacité du tri, identifier les défis potentiels et évaluer l'impact des initiatives de recyclage mises en place par la société. Cela permet également de mettre en lumière la gestion des déchets qui peuvent informer des stratégies futures pour l'amélioration continue des processus de tri et de recyclage.

Février 2022:

De la même façon ici un échantillon représentatif de 1242,51 kg a été sélectionné dans le flux total de déchets qui pesait 1 986.460 Tonnes.



Répartition de la masse de l'échantillon (Fev_22)

Figure 9

L'échantillon de 1242.51 kg représente environ 0.06% de la masse totale des déchets de février, offrant ainsi un aperçu représentatif de la composition des déchets traités durant ce mois. Cet échantillonnage aide à comprendre les proportions des objets valorisables et non valorisables, et constitue une base pour l'amélioration des processus de tri. Il permet également d'identifier les points d'amélioration potentiels, tels que la réduction du taux de "Refus", qui s'élève à 10.3% de l'échantillon, et suggère une marge d'optimisation pour la récupération et la valorisation des déchets.

V Amélioration des Techniques d'Échantillonnage pour une Évaluation Fiable du Tri des Déchets

Un protocole d'échantillonnage mal adapté peut entraîner une évaluation erronée de la qualité du tri, résultant en une sous-évaluation ou une surévaluation qui risque d'influencer négativement les décisions stratégiques et l'efficacité du système de recyclage. L'enjeu est donc de mettre en place une méthodologie d'échantillonnage qui allie précision des résultats et rentabilité.

Dans cette section, nous abordons les fondements de l'échantillonnage comme outil d'analyse statistique, les défis de son implémentation dans le contexte du tri des déchets, ainsi que les approches envisageables pour relever ces défis. L'objectif est de renforcer la justesse des évaluations et d'optimiser la performance du processus de tri

Pour aborder la question posée, nous nous baserons sur une étude préliminaire qui a modélisé la situation de manière simplifiée (Modélisation jouet du problème) ici . Nous reprendrons les mêmes données et échantillons précédant pour les appliquer et évaluer notre approche qui va nous permettre de donner un critère pour choisir l'échantillon par rapport à une taille de population données.

VI Simulation

Dans le contexte de notre modèle simplifié du problème, nous avons précédemment tenté de trier trois objets distincts et de calculer leurs estimateurs ainsi que les intervalles de confiance [Modélisation jouet du problème]. Cette méthode a été testée et validée. Nous envisageons maintenant d'étendre cette approche à douze objets différents, en utilisant des données réelles issues des essais réalisés au centre de tri d'Angers.

Nous conserverons les mêmes notations, définitions et paramètres que ceux de la modélisation simplifiée du problème. Nous vous encourageons donc vivement à consulter cette modélisation pour bien comprendre chaque définition des paramètres. Θ_{ij} représente la proportion de masse d'objets de type i présents dans le bac j tel que :

$$\Theta_{ij} = \frac{\text{"Masse réelle d'objets de type } i \text{ dans le bac } j\text{"}}{\text{Masse totale des objets dans le bac } j}$$

En supposant les variables X_i et Y_i qui représentent respectivement la masse et le type d'objet i, dans ce cas on a :

$$\Theta_{ij} = \frac{\mathbb{E}(X_i \cdot \mathbb{1}_{\{Y_i = j\}})}{\mathbb{E}(X_i)}$$

Dans cette perspective que l'estimateur $\hat{\Theta}_{ij}$, qui mesure la proportion de masse estimée d'objets de type i dans le bac j, suit une distribution qui peut être approchée par une loi normale multivariate d'aprés notre modèle jouet : [ici]

$$\hat{\Theta}_{ij} = \frac{\mu_i p_i + \mu Z_j}{\mu (1 + Z_1 + \dots + Z_{12})} \text{ et } Z_j = \frac{1}{n} \sum_{i=1}^n \frac{(\mu_i \cdot \mathbb{1}_{Y_i = j} - \mu_j p_j)}{\mu},$$

 \sqrt{n} multiplié par Z_j suit asymptotiquement une loi normale tel que : $\sqrt{n}Z_j \to \mathcal{N}(0, \sigma_j^2)$ et σ_j^2 représente la variance de $\sqrt{n}Z_j$ et le vecteur $\mathbf{Z} = (Z_1, Z_2, \dots, Z_{12})$ converge aussi asymptotiquement vers la loi normale :

$$\mathbf{Z} \xrightarrow[n \to \infty]{} \frac{1}{\sqrt{n}} \mathcal{N} \left(\mathbf{0}, \frac{1}{\mu^2} \begin{pmatrix} p_1 \mu_1^2 & -p_1 p_2 \mu_1 \mu_2 & \cdots & -p_1 p_{12} \mu_1 \mu_{12} \\ -p_2 p_1 \mu_2 \mu_1 & p_2 \mu_2^2 & \cdots & -p_2 p_{12} \mu_2 \mu_{12} \\ \vdots & \vdots & \ddots & \vdots \\ -p_{12} p_1 \mu_{12} \mu_1 & -p_{12} p_2 \mu_{12} \mu_2 & \cdots & p_{12} \mu_{12}^2 \end{pmatrix} \right)$$

VI.1 Résultats Numériques

Pour Décembre 2021 :

Dans cette section, nous allons exploiter les données recueillies lors des essais menés au centre de tri d'Angers. Ces données seront appliquées à notre modèle généraliser à douze type d'objets, conçu pour simuler le problème de tri des déchets. L'objectif est d'observer comment les variations de la taille de l'échantillon affectent les intervalles de confiance ainsi que les taux de confiance associés.

Avec dix simulations pour tester ce modèle amélioré, nous avons pu affiner notre compréhension de la qualité du tri, ainsi nous observons les differents intervalle de confiances.

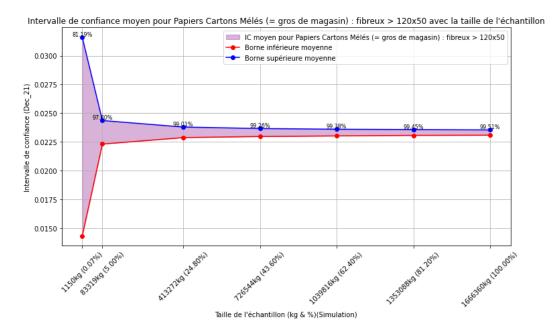


FIGURE 10 – Taux de confiance pour les Papiers en fonction de la taille de l'échantillon.

Ce graphique représente l'évolution de l'intervalle de confiance moyen pour les papiers cartons en fonction de la taille de l'échantillon. On observe qu'à mesure que la taille de l'échantillon augmente, l'intervalle de confiance moyen devient plus étroit, ce qui indique une estimation plus précise. Le taux de confiance pour les plus petits échantillons commence à 81.19% et s'accroît progressivement pour atteindre et se stabiliser autour de 99% lorsque la taille de l'échantillon s'élargit, suggérant une forte probabilité que la moyenne réelle soit capturée par cet intervalle. Cela suggère que des échantillons plus importants améliorent la fiabilité des estimations du tri de l'acier.

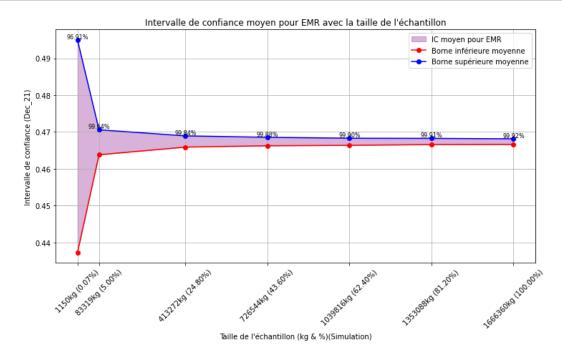


FIGURE 11 – Taux de confiance pour l'EMR en fonction de la taille de l'échantillon.

Ce graphique illustre l'intervalle de confiance moyen pour les Emballages Ménagers Recyclables (EMR) selon différentes tailles d'échantillon. Initialement, avec un très petit échantillon, l'intervalle de confiance est assez large, comme le montre la distance entre la borne inférieure et la borne supérieure moyenne, ce qui indique une estimation moins précise. Cependant, dès que la taille de l'échantillon augmente, l'intervalle se resserre significativement, indiquant une amélioration de la précision de l'estimation.

Le taux de confiance commence à un niveau élevé de 96.91% même pour le plus petit échantillon et s'accroît légèrement pour atteindre au-dessus de 99% avec l'augmentation de la taille de l'échantillon. Cette tendance indique que plus la taille de l'échantillon est grande, plus la probabilité est forte que l'intervalle de confiance contienne la véritable proportion de EMR correctement triés. Les résultats se stabilisent avec des taux de confiance dépassant 99,8%, ce qui reflète une grande confiance dans les estimations obtenues pour les plus grands échantillons analysés.

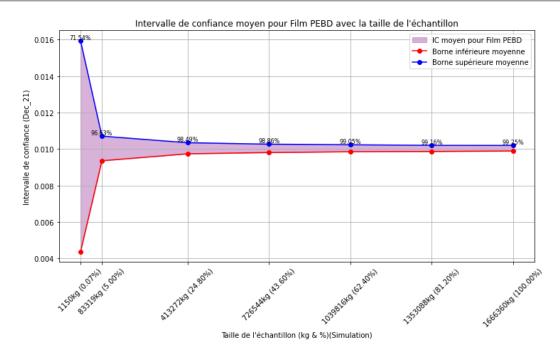


FIGURE 12 – Taux de confiance pour le Film PEBD, en fonction de la taille de l'échantillon.

Ce graphique représente l'intervalle de confiance moyen pour le film PEBD (Polyéthylène Basse Densité) en corrélation avec la taille de l'échantillon. La borne inférieure et la borne supérieure moyenne montrent qu'avec le plus petit échantillon, il y a une large variation, signifiant une précision moins fiable dans l'estimation. Cependant, cet intervalle se resserre rapidement à mesure que la taille de l'échantillon augmente, indiquant que la précision des estimations s'améliore.

Le taux de confiance initial pour le plus petit échantillon est de 77.54%, ce qui est relativement bas par rapport aux autres matériaux mais augmente rapidement pour dépasser les 96% même avec de faibles augmentations de la taille de l'échantillon. Il se stabilise audessus de 99,25% avec les échantillons plus importants, ce qui reflète une forte confiance dans l'exactitude des estimations pour le tri du film PEBD.

Il est important de noter que le film PEBD, en raison de sa très faible masse moyenne, peut être sujet à une plus grande variabilité dans le tri, ce qui peut expliquer le taux de confiance initial plus bas et la nécessité d'un échantillon légèrement plus grand pour atteindre des niveaux de confiance comparables à ceux d'autres matériaux plus lourds.

Les autres graphiques, qui seront placés en annexe, fourniront des détails supplémentaires pour chaque type de matériau individuel.

Nous allons maintenant proposer un critère global pour évaluer la variation de la taille de l'échantillon et du niveau de confiance en fonction du volume des déchets à trier.

Entre 0.07 % à 1% de la taille de l'échantillon :

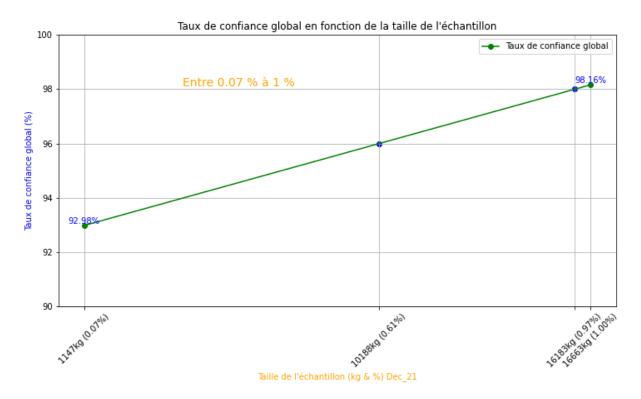


FIGURE 13 – Taux de confiance Globale en fonction de la taille de l'échantillon.

On peut observer une progression claire et linéaire du taux de confiance à mesure que la taille de l'échantillon augmente, partant de 92.98% pour un échantillon très petit à environ 0.07% de la masse totale, jusqu'à atteindre près de 98.16% pour un échantillon représentant 1% de la masse totale de l'échantillon.

Cette tendance suggère que l'accroissement de la taille de l'échantillon contribue significativement à améliorer la précision des estimations de tri. Elle montre également que des taux de confiance élevés sont réalisables avec des augmentations proportionnelles de la taille de l'échantillon.

Entre 1.0 % à 10 % de la taille de l'échantillon :

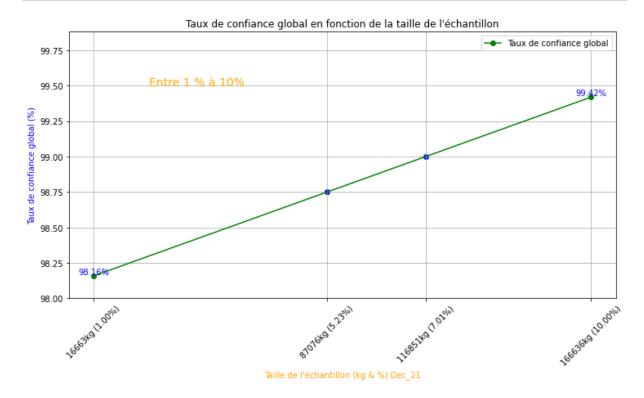


FIGURE 14 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Commençant de 98.16% pour un échantillon relativement petit (représentant 1.0% de la masse totale) et culminant à près de 99.42% pour un échantillon plus conséquent (10% de la masse totale). Cette tendance linéaire aussi illustre que l'augmentation de la taille de l'échantillon renforce de manière significative la fiabilité des estimations statistiques obtenues. Elle met également en lumière que pour atteindre un niveau de confiance extrêmement élevé, l'accroissement de la taille de l'échantillon est essentiel, soulignant l'importance de choisir des tailles d'échantillon appropriées pour garantir la validité et l'utilité des analyses de données dans la gestion et l'évaluation des processus de tri des déchets. (voir annexe pour les autres Figures)

Pour Janvier 2022:

Nous avons utilisé des échantillons de tailles variées, qui représentent différentes proportions de la masse totale des déchets traités pour chaque catégorie.

Ces graphiques présentent les intervalles de confiance moyens pour chaque type de déchet en fonction de la taille de l'échantillon.

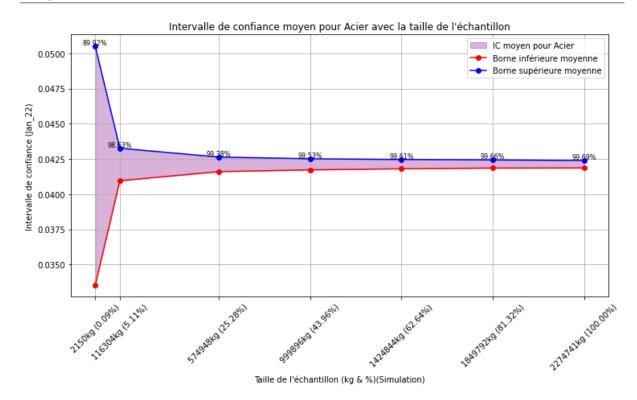


FIGURE 15 – Taux de confiance pour l'acier en fonction de la taille de l'échantillon.

Après dix simulations ont révélé un taux de confiance moyen de 89.92% de tri correct de l'acier, avec un échantillon de 0,09 % par rapport au total traité en **Janvier 2022**. Ces résultats démontrent l'efficacité de notre approche généralisée et l'importance de la taille de l'échantillon dans l'évaluation précise de la qualité du tri des déchets.

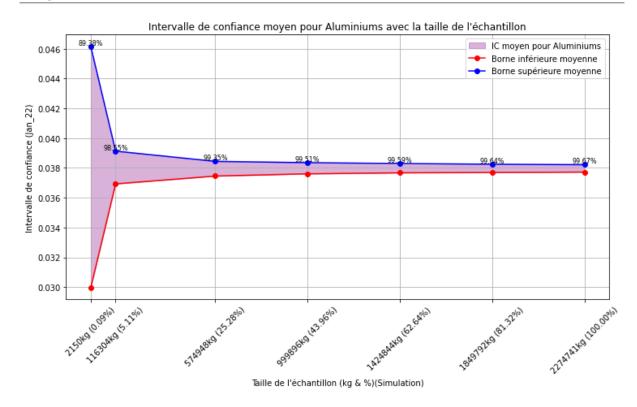


Figure 16

Ce graphique pour l'aluminium révèle que la précision des estimations qui s'améliore avec l'augmentation de la taille de l'échantillon. Un échantillon représentant 0,09% de la totalité des déchets a fourni un taux de confiance significative, qui se stabilise et se resserre à mesure que la taille de l'échantillon augmente, suggérant une estimation plus précise pour des échantillons plus importants. Cela se manifeste par un taux de confiance qui commence autour de 89,38% et qui atteint jusqu'à 98,65% pour une taille d'échantillons de 5.11%.

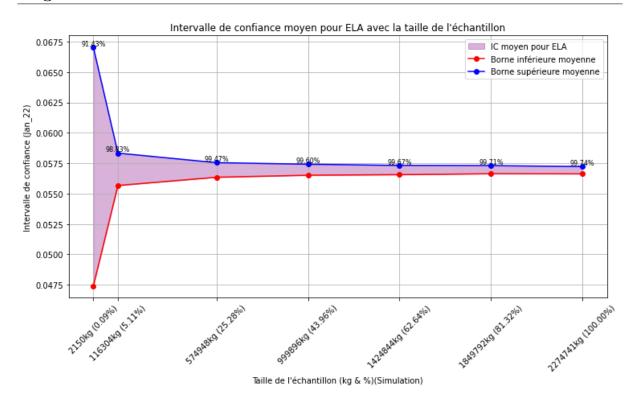


FIGURE 17

Le graphique concernant l'ELA (Briques alimentaires) indique une tendance similaire à celle observée pour l'aluminium, où la fiabilité de l'estimation augmente avec la taille de l'échantillon. Partant d'un intervalle de confiance plus large pour les plus petits échantillons, qui témoigne d'une plus grande incertitude, la précision s'améliore nettement à mesure que l'échantillon s'agrandit. Les données démontrent une confiance initiale de 91.43% pour les échantillons les plus petits, qui progresse à 98% pour les échantillons de 5%, affirmant l'exactitude accrue de l'évaluation du tri de l'ELA avec des échantillons plus représentatifs.

On constate que la taille de l'échantillon compris entre 0.05% à 5% de la masse totale des déchets nous avons dépasser largement le taux de confiance souhaiter ainsi, cela confirme que pour garantir des conclusions statistiques fiables, il est essentiel de choisir une taille d'échantillon adéquate.

Entre 0.05 % à 1 % de la taille d'échantillon :

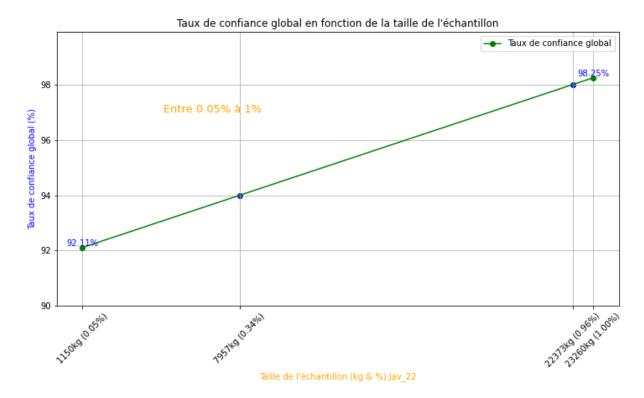


FIGURE 18 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Le taux de confiance global commence à 92% pour une très petite taille d'échantillon et augmente progressivement pour atteindre pratiquement plus de 98% à mesure que la taille de l'échantillon augmente jusqu'à 1.0% de la masse totale des déchets.

Entre 1 % à 10% de la taille de l'échantillon :

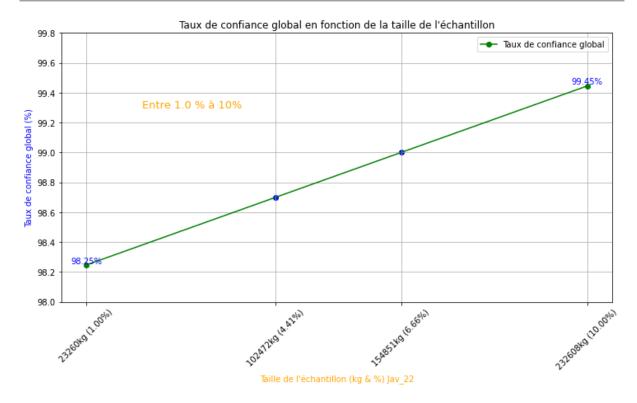


FIGURE 19 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Dans une plage de taille d'échantillon allant de 2% à 10%, le taux de confiance démarre à environ 98.73% et monte à près de 99.43%, montrant une tendance similaire à une amélioration de la fiabilité avec des échantillons plus grands.

Pour Février 2022 :

En appliquant la même méthodologie d'analyse que celle utilisée pour les données de décembre et janvier, les résultats pour un autre ensemble de données montrent une progression similaire dans le taux de confiance global en fonction de la taille de l'échantillon.

Pour ce nouvel ensemble de données, nous observons un saut initial significatif dans le taux de confiance global à mesure que la taille de l'échantillon passe de très petite à modeste. Le taux de confiance augmente rapidement, atteignant un plateau de stabilité à un niveau élevé pour des tailles d'échantillon plus importantes, ce qui indique que la taille de l'échantillon optimale pour une estimation précise a été atteinte et que des échantillons plus grands n'augmentent pas de manière significative la confiance dans les estimations du taux de tri correct.

Entre 0.06~% à 1~% de la taille de l'échantillon :

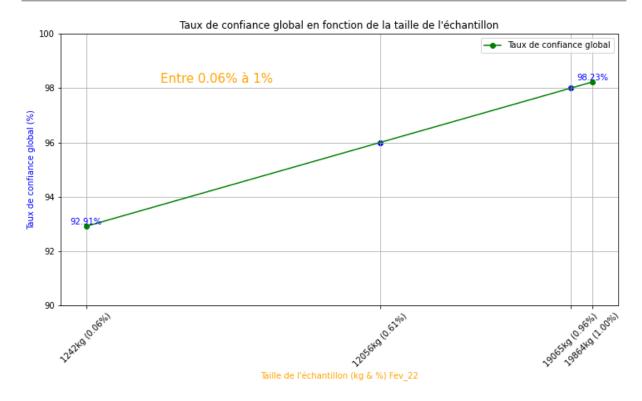


FIGURE 20 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Débutant à 91.93% pour un échantillon représentant seulement 0.06% de la masse totale, le taux de confiance augmente rapidement et atteint 98.35% pour un échantillon de 1.5%. Cela indique qu'une taille d'échantillon accrue dans cette gamme a un impact significatif sur la fiabilité des résultats.

Entre 1 % à 10% de la taille de l'échantillon :

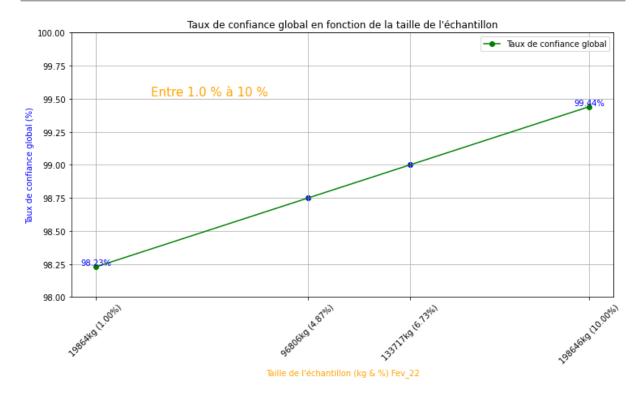


FIGURE 21 – Taux de confiance Globale en fonction de la taille de l'échantillon.

VI.2 Analyse des Taux de Confiance par Catégorie Maj. et Min.

Jusqu'à présent, nous avons observé qu'il existe des produits majoritaires et des produits minoritaires dans la répartition de notre échantillon choisi. Cela explique pourquoi nous ne pouvons pas obtenir les mêmes taux de confiance pour un échantillon donné. C'est pourquoi nous allons représenter les produits ou les déchets les mieux caractérisés ainsi que ceux qui sont plus difficiles à caractériser. Prenons l'exemple pour Dec. 2021

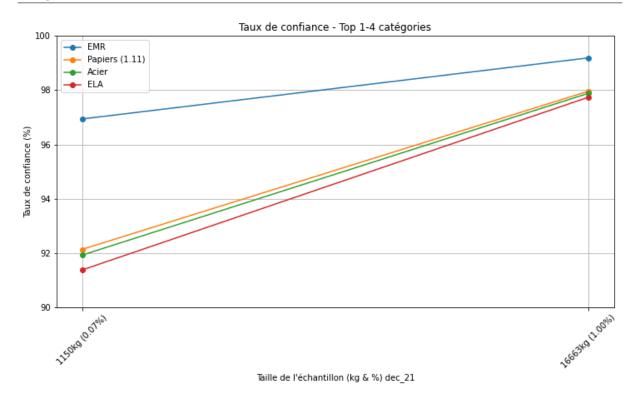


FIGURE 22 – Taux de confiance par Catégories 1 à 4

Ce graphique montre que les taux de confiance pour EMR, Papiers (1.11), Acier et ELA entre 0.07~% à 1.0~% des tailles d'échantillon.

- EMR: Cette catégorie affiche le taux de confiance le plus élevé, commençant à environ 97 % et atteignant environ 99 %. Il y a une augmentation constante, indiquant un niveau élevé de fiabilité dans cette catégorie à travers différentes tailles d'échantillon.
- Papiers (1.11): Le taux de confiance commence avec plus 92 % et atteint près de 98 %. Suivi proche des autres catégories sauf EMR, montrant une fiabilité significative.
- Acier : Commence légèrement au-dessus de 92 % et se termine près de 99 %. Montre une cohérence similaire aux Papiers, indiquant une confiance fiable dans cette catégorie.
- ELA: Commence juste en dessous de 92 % et se termine près de 97 %. Tendance similaire à l'Acier, démontrant une confiance fiable à travers les tailles d'échantillon.

Les catégories affichent des taux de confiance élevés, ce qui montrent que les déchets majoritaires (par exemple, JRM , EMR) sont généralement mieux caractérisés que les déchets minoritaires (par exemple, Film PEBD, Aluminium Souple). Cela s'explique par plusieurs facteurs.

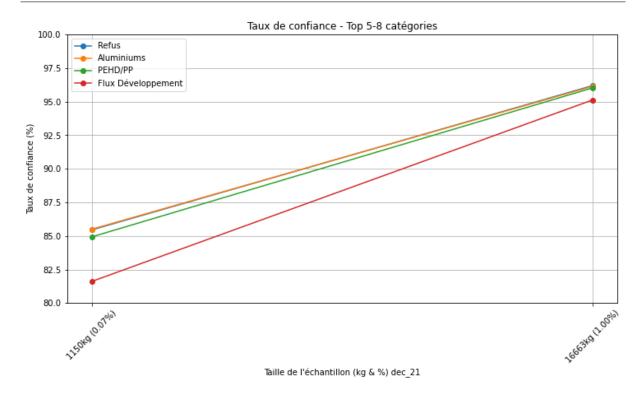


FIGURE 23 – Taux de confiance par Catégories 5 à 8

Ici ce graphique montre que les taux de confiance pour les catégories 5 à 8 : Refus, Aluminiums, PEHD/PP, et Flux Développement.

- Refus: Commence à environ 85 % et augmente jusqu'à environ 96 %. Indique une amélioration significative du taux de confiance avec des tailles d'échantillon plus grandes.
- Aluminiums : Le taux de confiance initial est légèrement au-dessus de 87 %, augmentant jusqu'à environ 87 %. Une fiabilité élevée est maintenue, montrant un suivi proche avec le PEHD/PP.
- **PEHD/PP**: Commence à environ 85 % et atteint presque 97 %. Cohérent avec les Aluminiums, indiquant des niveaux de confiance élevés.
- Flux Développement : Commence à environ 81 % et se termine près de 95 %. Montre l'amélioration la plus significative, reflétant une augmentation substantielle de la confiance avec des tailles d'échantillon plus grandes.

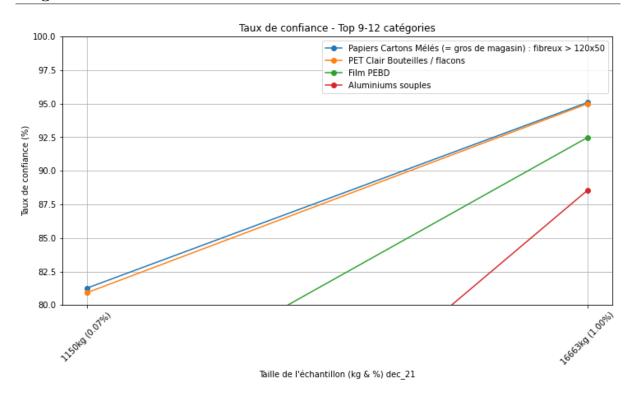


FIGURE 24 – Taux de confiance par Catégories 9 à 12

Les catégories avec des taux de confiance plus faibles nécessitent souvent des échantillons plus importants pour obtenir des estimations précises. Les difficultés de tri pour certaines catégories comme le Film et Aluminium souples peuvent être liées à la présence d'objets minoritaires qui sont difficiles à identifier ou à séparer. L'analyse des données de tri permet de comprendre les défis du processus et d'identifier les points d'amélioration pour optimiser l'efficacité du tri et la qualité des produits recyclés. Ainsi proposer un plan d'échantillonnage comme échantillonnage par stratification. Le principe de la stratification : découper la population en sous-ensembles appelés strates et réaliser un sondage dans chacune d'elles. Avec plus de détaille (ici Chapitre 3 [1]). On procède de la même façons pour les données de Janvier et Février, et nous allons juste laisser les graphique en annexe pour confirmer l'analogie de ces différents catégories.

En conclusion...

En conclusion, cette étude met en évidence l'importance cruciale de la taille de l'échantillon dans l'évaluation précise de la qualité du tri des déchets. Les résultats obtenus montrent clairement que des échantillons plus grands augmentent significativement la fiabilité des estimations statistiques, comme le démontre l'amélioration progressive des taux de confiance avec l'augmentation de la taille des échantillons.

En particulier, nos résultats soulignent l'importance d'utiliser des taille d'échantillon appropriées pour garantir la validité des analyse de tri des déchets, Par exemple pour garantir un taux de confiance de 98%, il est nécessaire de choisir des échantillons représentant entre 0.85% et 1% de la masse totale de déchets, correspondant à une masse de 14 à 16 tonnes par mois parmi 1666 Tonnes.

Ces observations confirment que pour garantir des conclusions statistiques fiables, il est essentiel de choisir des tailles d'échantillon adéquates. Cela permet non seulement d'augmenter la précision des estimations mais également de renforcer la validité et l'utilité des analyses de données dans la gestion et l'évaluation des processus de tri des déchets

En somme, cette étude souligne l'importance de la méthodologie d'échantillonnage et des analyses statistiques rigoureuses pour améliorer la gestion des déchets et optimiser les pratiques de tri, contribuant ainsi à une meilleure efficacité environnementale et opérationnelle des systèmes de gestion des déchets.

Bibliographie

- [1] Laurent Rouvière, Introduction aux sondages
- [2] TRIVALIS: Manuel Flux entrants Version Janvier 2020 et d'autre documents comme Cahier des garanties de performances, sont disponible à l'adresse: https://drive.google.com/drive/folders/1-79q8KGaKOGEbJcniE8GNMCcB9YGxSAO?usp=drive_link
- [3] Íneris : Echantillonnage des déchets solides Nombre de particules d'échantillons représentatifs en l'absence d'hypothèses de distribution, disponible à l'adresse : https://www.ineris.fr/sites/ineris.fr/files/contribution/Documents/Rapport-20-200356-1477888_DRC-01_Echantillonnage%20v2.pdf
- [4] Vauché, Historique de l'Etablissement disponible à l'adresse :https://www.racine-d-ardennes.fr/histoire-des-etablissements-vauche-gerald-dardart-c2x3382276

VII Annexe

VII.1 Modélisation jouet du problème

Pour simplifier le travail, les données seront réduites à trois types de déchets, de sorte que l'échantillon contient trois types présentant des proportions p_1, p_2 , et p_3 , avec $p_1+p_2+p_3=1$. Ces types sont triés dans trois bacs (bac1, bac2, bac3) par trois machines, illustrées dans la figure ci-contre.

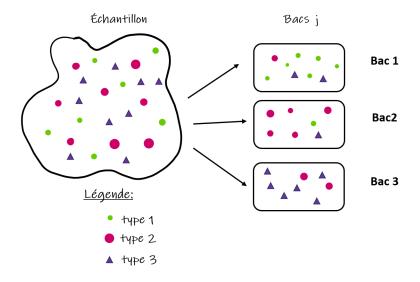


FIGURE 25 – Illustration du problème

La figure illustre un processus de tri d'objets basé sur leur type, suggérant une analogie avec des matériaux recyclables tels que le carton, le plastique, le verre, etc. Dans cette représentation :

- Les cercles verts représentent des objets d'un type spécifique, comme du verre, qui peuvent varier en taille et en poids.
- Les cercles roses symbolisent un autre type d'objet, tels que des plastiques, également variés en taille et en poids.
- Les triangles violets correspondent à un troisième type d'objet, par exemple du carton ou du papier, qui diffèrent aussi en taille et en poids.

Les objets sont répartis en trois bacs distincts, suggérant une séparation des matériaux pour un traitement spécifique dans une installation de recyclage. Cette séparation est cruciale pour maintenir l'efficacité du processus de recyclage et minimiser la contamination croisée entre différents types de matériaux recyclables.

VII.2 Déclaration des paramètres pour le cas jouet

Dans cette partie on va introduire les différents paramètres que on aura besoin pour le modèle jouet.

Soit M une matrice de tri que on peut contrôler entre les déchets à trier vers les différents bacs et une matrice de résultat Θ qu'on observe à la fin tels que :

$$M = \begin{pmatrix} m_1 & m_2 & m_3 \end{pmatrix}$$

$$\Theta = \begin{pmatrix} \Theta_{11} & \Theta_{12} & \Theta_{13} \\ \Theta_{21} & \Theta_{22} & \Theta_{23} \\ \Theta_{31} & \Theta_{32} & \Theta_{33} \end{pmatrix} \leftarrow \begin{pmatrix} \text{objet i} \\ \\ \\ \\ \\ \end{pmatrix}$$

- \bullet m_i : représente la masse moyenne unitaire d'objets de type i
- Θ_{ij} : est la proportion de nombre d'objet du type i dans le bac j (qu'on va essayer à estimer)
- $\bullet \ n$: est le nombre total d'objets qu'on observe dans l'échantillon
- N_{ij} : est le nombre d'objets du type i dans le bac j
- p_i : représente la proportion d'avoir un objet dans le bac i $p_i = \frac{1}{3}$

Dans le cas idéal où le tri est parfait, toutes les valeurs hors de la diagonale de M et Θ seraient nulles, et les valeurs sur la diagonale seraient égales à 1, signifiant que tous les objets de type i se retrouvent exclusivement dans le bac i et dans aucune autre. Ce cas n'est pas réalisable d'où on cherche à avoir des matrices avec des valeurs proches de 1 sur les diagonales.

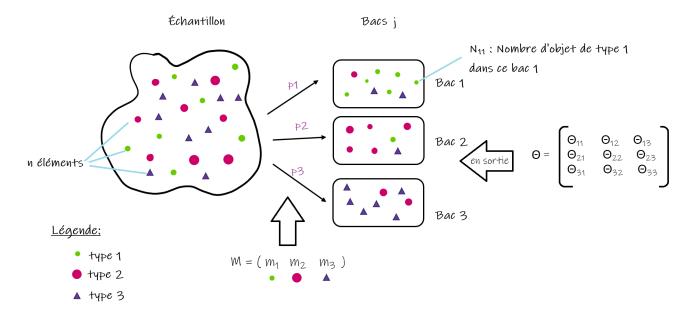


FIGURE 26 – Figure explicative de l'échantillonnage et des paramètres

VII.3 Estimateur des paramètres

Dans le modèle Θ_{ij} représente la proportion réelle du nombre d'objets de type i présents dans le bac j. Cette proportion est définie comme le rapport entre le nombre d'objets de type i dans un bac spécifique j et le nombre total d'objets dans ce même bac j, toutes catégories confondues. Si on avait les données complètes, l'expression de Θ_{ij} pourrait être écrite comme suit :

$$\Theta_{ij} = \frac{\text{nombre r\'eel d'objets de type } i \text{ dans le bac } j}{\text{nombre total d'objets dans le bac } j}$$

avec $\sum_{j=1}^{3} \Theta_{ij} = 1$ vu que Θ_{ij} représente la part des objets de type i qui sont correctement triés dans le bac j, et cette valeur est normalisé pour tenir en compte la répartition totale de chaque type d'objet dans l'ensemble du système de tri.

Dans le contexte où N_{ij} est défini comme le nombre d'objets du type i dans le bac j tel que :

$$\sum_{i=1}^{3} \sum_{j=1}^{3} N_{ij} = n$$

On peut donc utilisée ces expressions pour des calculs des estimations de proportions. Voici comment elle peut être employée pour calculer l'estimateur de la proportion $\hat{\Theta}_{ij}$

L'estimateur $\hat{\Theta}_{ij}$ représente la proportion estimée du type d'objet i dans le bac j par rapport à la somme totale des objets dans ce bac. L'expression pour $\hat{\Theta}_{ij}$ en utilisant les N_{ij} est :

$$\hat{\Theta}_{ij} = \frac{N_{ij}}{n}$$

où:

- N_{ij} est le nombre d'objets de type i dans le bac j
- \bullet n nombre total d'objets échantillonnés.

Cette formule calcule la part relative des objets de type i parmi tous les objets dans le bac j, ce qui est une information utile pour comprendre comment les déchets sont distribués entre différents bacs et pour optimiser le processus de tri.

VII.4 Loi de l'estimateur du paramètre

Etape 1

Comme l'estimateur de la proportion $\hat{\Theta}_{ij} = \frac{N_{ij}}{n}$ est la proportion de succès (nombre d'objets de type i dans le bac j) dans n essais (nombre d'échantillons).

Pour connaître la loi de $\hat{\Theta}_{ij}$, il suffit de savoir la loi de N_{ij} . Pour cela on va définir \mathcal{O}_i qui indique un objet i dans l'échantillon et \mathcal{T}_i le type de cet objet.

 N_{ij} peut être écrit ainsi en indiquant que l'objet de type i va au bac j :

$$N_{ij} = \sum_{k=1}^{n} \mathbb{1}_{\{\mathcal{O}_k = j\}} \mathbb{1}_{\{\mathcal{T}_k = i\}}$$

En remarquant que $\mathbb{1}_{\{\mathcal{O}_k=j\}}\mathbb{1}_{\{\mathcal{T}_k=i\}}$ suit une loi de Bernoulli $\mathcal{B}(\mathbb{P}(\mathcal{O}=j,\mathcal{T}=i))$ et comme N_{ij} est la somme de $\mathbb{1}_{\{\mathcal{O}_k=j\}}\mathbb{1}_{\{\mathcal{T}_k=i\}}$ sur n donc N_{ij} suit une loi binomiale $\mathcal{B}(n,\mathbb{P}(\mathcal{O}=j,\mathcal{T}=i))$.

 N_{ij} est la somme d'une suite de variables aléatoires réelles définies sur le même espace de probabilité, indépendantes et identiquement distribués suivant la même loi (loi de Bernoulli). D'après le théorème centrale limite, on aura

$$\sqrt{n}\left(\hat{\Theta}_{ij} - \Theta_{ij}\right) \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}\left(0, \Theta_{ij}(1 - \Theta_{ij})\right)$$

Approximation Normale

Pour un grand nombre d'objets (n grand), en vertu du théorème central limite, l'estimateur $\hat{\Theta}_{ij}$ peut être approché par une distribution normale si certaines conditions sont remplies. La moyenne et la variance de l'estimateur $\hat{\Theta}_{ij}$ dans cette approximation normale sont :

• Moyenne $(\mu): \Theta_{ij}$

• Variance (σ^2) : $\frac{\Theta_{ij}(1-\Theta_{ij})}{n}$.

Ainsi, la loi normale approchée pour $\hat{\Theta}_{ij}$ est :

$$\hat{\Theta}_{ij} \sim \mathcal{N}\left(\Theta_{ij}, \frac{\Theta_{ij}(1 - \Theta_{ij})}{n}\right)$$

où \mathcal{N} désigne la distribution normale.

Expression de la loi : Sous l'approximation normale, l'expression pour la distribution de $\hat{\Theta}_{ij}$ serait :

$$\hat{\Theta}_{ij} \approx \mathcal{N}\left(\frac{N_{ij}}{n}, \frac{\frac{N_{ij}}{n}\left(1 - \frac{N_{ij}}{n}\right)}{n}\right)$$

Il est crucial de noter que cette approximation normale est plus précise lorsque le nombre total d'objets dans le bac j est suffisamment grand et que les conditions pour le théorème central limite sont remplies. Pour des nombres plus petits, la distribution exacte binomiale doit être utilisée pour des calculs précis.

Etape 2

Dans le cas où on introduit la masse on obtient :

Dans cette partie on va introduire les variables X_i et Y_i qui représente respectivement la masse et le type d'objet i, dans ce cas on a :

$$M_1 = \sum_i X_i \cdot \mathbb{1}_{Y_i = 1}$$

 M_1 : est la masse d'objets de type 1 sur la ligne de tri.

 X_i : représente la masse de l'objet i si cet objet est de type 1 et se trouve dans le bac 1. $\mathbb{1}_{Y_i=1}$: est une fonction indicatrice qui vaut 1 si l'objet i est de type 1, et 0 sinon.

$$\Theta_{11} = \frac{\mathbb{E}(X_i \cdot \mathbb{1}_{Y_i = 1})}{\mathbb{E}(X_i)}$$

Dans ce cas l'estimateur de paramètre devient :

$$\hat{\Theta}_{11} = \frac{M_1}{\sum_i X_i}$$

C'est l'estimateur de la proportion de déchets de type 1 dans le bac 1 qui permet d'évaluer l'efficacité avec laquelle les objets de type 1 sont triés et déposés dans le bac 1. Cela aide à identifier si les objets sont correctement triés selon leur type.

Ainsi,

$$\hat{\Theta}_{11} = \frac{\frac{1}{n} \sum_{i=1}^{n} X_i \cdot \mathbb{1}_{Y_i = 1}}{\frac{1}{n} \sum_{i=1}^{n} X_i}$$

Soit μ_1 : la moyenne des masses des objets sachant qu'ils sont de type 1 et μ : serait la moyenne pondérée des masses pour tous les types d'objets dans le système, où chaque masse moyenne μ_i est pondérée par la probabilité p_i de l'objet de type i.

$$\mu_1 = \mathbb{E}(X_i | Y_i = 1)...$$
 et $p_1 = \mathbb{P}(Y_i = 1)$ tel que $\mu = p_1 \mu_1 + p_2 \mu_2 + p_3 \mu_3$

Ce qui implique que :

$$\hat{\Theta}_{11} = \frac{\mu_1 p_1 + \frac{1}{n} \sum_{i=1}^{n} (\mu_i \cdot \mathbb{1}_{Y_i = 1} - \mu_1 p_1)}{\mu_1 p_1 + \mu_2 p_2 + \mu_3 p_3 + \frac{1}{n} \gamma}$$

Avec
$$\gamma = \sum_{i=1}^{n} (\mu_i \cdot \mathbb{1}_{Y_i=1} - \mu_1 p_1) + \sum_{i=1}^{n} (\mu_i \cdot \mathbb{1}_{Y_i=2} - \mu_2 p_2) + \sum_{i=1}^{n} (\mu_i \cdot \mathbb{1}_{Y_i=3} - \mu_3 p_3)$$

En posant:

$$Z_1 = \frac{1}{n} \sum_{i=1}^{n} \frac{(\mu_i \cdot \mathbb{1}_{Y_i=1} - \mu_1 p_1)}{\mu},$$

$$Z_2 = \frac{1}{n} \sum_{i=1}^{n} \frac{(\mu_i \cdot \mathbb{1}_{Y_i=2} - \mu_2 p_2)}{\mu},$$

et

$$Z_3 = \frac{1}{n} \sum_{i=1}^{n} \frac{(\mu_i \cdot \mathbb{1}_{Y_i=3} - \mu_3 p_3)}{\mu},$$

Donc on obtient:

$$\hat{\Theta}_{11} = \frac{\mu_1 p_1 + \mu Z_1}{\mu (1 + Z_1 + Z_2 + Z_3)}$$

 \sqrt{n} multiplié par Z_1 suit asymptotiquement une loi normale tel que : $\sqrt{n}Z_1 \to \mathcal{N}(0, \sigma_1^2)$ et σ_1^2 représente la variance de $\sqrt{n}Z_1$ qui peut être approximée par :

$$\operatorname{Var}(\sqrt{n}Z_1) = n \cdot \operatorname{Var}\left(\frac{1}{n} \sum_{i=1}^{n} \frac{(\mu_i \cdot \mathbb{1}_{Y_i=1} - \mu_1 p_1)}{\mu}\right) = n \cdot \operatorname{Var}(Z_1)$$

En développant la variance de Z_1 on obtient :

$$Var(Z_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n} \left(\frac{\mu_i \cdot 1_{\{Y_i=1\}}}{\mu}\right) - \frac{\mu_1 p_1}{\mu}\right) = \frac{1}{\mu^2 n^2} Var\left(\sum_{i=1}^{n} \mu_i \cdot 1_{\{Y_i=1\}}\right)$$

En isolant les termes de variance de $\mu_i \cdot \mathbb{1}_{Y_i=1}$ ainsi que de $\mu_1 p_1$, on obtient :

$$\operatorname{Var}\left(\sum_{i=1}^{n} \mu_1 \cdot 1_{\{Y_i=1\}}\right) = \mu_1^2 n. \mathbb{P}(Y_i = 1) \text{ et } \operatorname{Var}(\mu_1 p_1) = 0$$

Par conséquent,

$$\operatorname{Var}(\sqrt{n}Z_1) = \frac{p_1\mu_1^2}{\mu^2}$$

Considérons un vecteur aléatoire $\mathbf{Z} = (Z_1, Z_2, Z_3)$ représentant des variables normalisées pour les proportions des types d'objets dans un bac de tri. La matrice de covariance Σ de \mathbf{Z} contient les covariances entre les composantes du vecteur, définies comme suit :

$$Cov(Z_i, Z_j) = \mathbb{E}[(Z_i - \mathbb{E}[Z_i])(Z_j - \mathbb{E}[Z_j])]$$

où $\mathbb{E}[Z_i]$ est l'espérance de Z_i .

La matrice de covariance pour \mathbf{Z} est donc :

$$\Sigma = \begin{pmatrix} \operatorname{Var}(Z_1) & \operatorname{Cov}(Z_1, Z_2) & \operatorname{Cov}(Z_1, Z_3) \\ \operatorname{Cov}(Z_2, Z_1) & \operatorname{Var}(Z_2) & \operatorname{Cov}(Z_2, Z_3) \\ \operatorname{Cov}(Z_3, Z_1) & \operatorname{Cov}(Z_3, Z_2) & \operatorname{Var}(Z_3) \end{pmatrix}$$

Où:

- $Var(Z_i)$ est la variance de Z_i ,
- $Cov(Z_i, Z_j)$ est la covariance entre Z_i et Z_j .

En se basant sur l'explication précédente, les variances de $\mathbb{Z}_1, \mathbb{Z}_2,$ et \mathbb{Z}_3 sont données par :

$$\operatorname{Var}(\sqrt{n}Z_1) = n.\operatorname{Var}(Z_1) = \frac{p_1\mu_1^2}{\mu^2}$$

Dans ce cas on peut aussi calculer la covariance de Z_1 Et Z_2 et puis en déduire les autres covariances.

En effet,

$$\begin{aligned} \operatorname{Cov}(\sqrt{n}Z_{1},\sqrt{n}Z_{2}) &= \mathbb{E}(\sqrt{n}Z_{1},\sqrt{n}Z_{2}) \\ &= \frac{n}{n^{2}\mu^{2}} \sum_{i,j=1}^{n} \mathbb{E}(\mu_{1} \cdot \mathbb{1}_{Y_{i}=1} - \mu_{1}p_{1})(\mu_{2} \cdot \mathbb{1}_{Y_{j}=2} - \mu_{2}p_{2}) \\ &= \frac{1}{\mu^{2}n^{2}} \sum_{k=1}^{n} \mathbb{E}(\mu_{1} \cdot \mathbb{1}_{Y_{k}=1} - \mu_{1}p_{1})(\mu_{2} \cdot \mathbb{1}_{Y_{k}=2} - \mu_{2}p_{2}) \\ &= \frac{1}{\mu^{2}} \mathbb{E}(\mu_{1} \cdot \mathbb{1}_{Y_{k}=1} - \mu_{1}p_{1})(\mu_{2} \cdot \mathbb{1}_{Y_{k}=2} - \mu_{2}p_{2}) \qquad (Y_{i} \perp \!\!\! \perp Y_{j} \text{ si } i \neq j) \\ &= \frac{1}{\mu^{2}} \left[\mathbb{E}(\mu_{1} \cdot \mathbb{1}_{Y_{k}=1} \cdot \mu_{2} \cdot \mathbb{1}_{Y_{k}=2}) - \mathbb{E}(\mu_{1} \cdot \mathbb{1}_{Y_{k}=1} \cdot \mu_{2}p_{2}) - \mathbb{E}(\mu_{1}p_{1} \cdot \mu_{2} \cdot \mathbb{1}_{Y_{k}=2}) \right. \\ &\quad + \mathbb{E}(\mu_{1}p_{1} \cdot \mu_{2}p_{2}) \right] \\ &= \frac{1}{\mu^{2}} \left(0 - \mu_{1}p_{1}\mu_{2}p_{2} - \mu_{1}p_{1}\mu_{2}p_{2} + \mu_{1}p_{1}\mu_{2}p_{2} \right) \\ &= \frac{-p_{1}p_{2}\mu_{1}\mu_{2}}{\mu^{2}} \end{aligned}$$

Donc

$$Cov(\sqrt{n}Z_1, \sqrt{n}Z_2) = \frac{-p_1p_2\mu_1\mu_2}{\mu^2}$$

Si les proportions des différents types d'objets dans les bacs de tri sont indépendantes, les éléments hors de la diagonale de Σ seraient nuls et les éléments diagonaux seraient les variances des Z_i .

Dans la réalité du tri des déchets, où certaines proportions sont susceptibles d'être corrélées, on doit estimer ces covariances à partir de données empiriques.

Le vecteur $\mathbf{Z} = (Z_1, Z_2, Z_3)$ Converge aussi asymptotiquement vers la loi Normal :

$$\mathbf{Z} \to \frac{1}{\sqrt{n}} \mathcal{N} \left(\mathbf{0}, \frac{1}{\mu^2} \begin{pmatrix} p_1 \mu_1^2 & -p_1 p_2 \mu_1 \mu_2 & -p_1 p_3 \mu_1 \mu_3 \\ -p_1 p_2 \mu_1 \mu_2 & p_2 \mu_2^2 & -p_2 p_3 \mu_2 \mu_3 \\ -p_1 p_3 \mu_1 \mu_3 & -p_2 p_3 \mu_2 \mu_3 & p_3 \mu_3^2 \end{pmatrix} \right), n \to \infty$$

Pour obtenir un intervalle de confiance fiable, il est nécessaire que la variance limite des données soit réduite. Ainsi, il est préférable que la masse (μ) de l'objet soit faible, car cela contribue à limiter la dispersion des observations autour de la moyenne. En d'autres termes, plus un objet est lourd, plus il sera difficile d'obtenir un intervalle de confiance étroit pour lui, car il sera observé moins souvent, ce qui augmente la variance des données.

L'estimateur $\hat{\Theta}_{ij}$, qui mesure la proportion estimée d'objets de type i dans le bac j, suit une distribution qui peut être approchée par une loi multinomiale dans le cas de données discrètes ou une loi normale sous certaines conditions en raison du théorème central limite.

$$\hat{\Theta}_{ij} = \frac{\mu_i p_i + \mu Z_j}{\mu (1 + Z_1 + Z_2 + Z_3)}$$

VII.5 Intervalle de Confiance

L'intervalle de confiance pour l'estimateur $\hat{\Theta}_{ij}$, qui représente la proportion estimée d'objets de type i dans le bac j, peut être dérivé en utilisant l'approximation normale lorsque le nombre d'objets est suffisamment grand pour que le théorème central limite s'applique :

$$\frac{\sqrt{n}(\hat{\Theta}_{ij} - \Theta_{ij})}{\sqrt{\Theta_{ij}(1 - \Theta_{ij})}} \xrightarrow{\mathcal{L} \text{ sous } \mathbb{P}} \mathcal{N}(0, 1) \text{ lorsque } n \to \infty.$$

Par définition de l'intervalle de confiance :

$$\mathbb{P}\left(-z_{\alpha/2} \le \frac{\sqrt{n}(\hat{\Theta}_{ij} - \Theta_{ij})}{\sqrt{\Theta_{ij}(1 - \Theta_{ij})}} \le z_{\alpha/2}\right) \approx 0.95.$$

Par la méthode de **Wald** on peut remplacer $\Theta_{ij}(1 - \Theta_{ij})$ par un estimateur $\hat{\Theta}_{ij}(1 - \hat{\Theta}_{ij})$ L'intervalle de confiance à $1 - \alpha$ pour $\hat{\Theta}_{ij}$ est donné par :

$$\hat{\Theta}_{ij} \pm z_{\alpha/2} \sqrt{\frac{\hat{\Theta}_{ij} (1 - \hat{\Theta}_{ij})}{n}}$$

où:

- $\hat{\Theta}_{ij}$ est l'estimateur de la proportion d'objets de type i dans le bac j,
- n est le nombre total d'objets dans le bac j,
- $z_{\alpha/2}$ est le quantile de la distribution normale standard correspondant au niveau de confiance souhaité 1α (pour un intervalle de confiance de 95%, $z_{\alpha/2} = 1.96$),
- $\sqrt{\frac{\hat{\Theta}_{ij}(1-\hat{\Theta}_{ij})}{n}}$ est l'erreur standard de l'estimateur $\hat{\Theta}_{ij}$.

Cet intervalle de confiance fournit une plage de valeurs dans laquelle on peut être raisonnablement sûrs que la vraie proportion d'objets de type i dans le bac j se trouve, en tenant compte du niveau de confiance choisi.

VII.6 Les graphes pour Décembre

Entre 10% à 30% de la taille de l'échantillon :

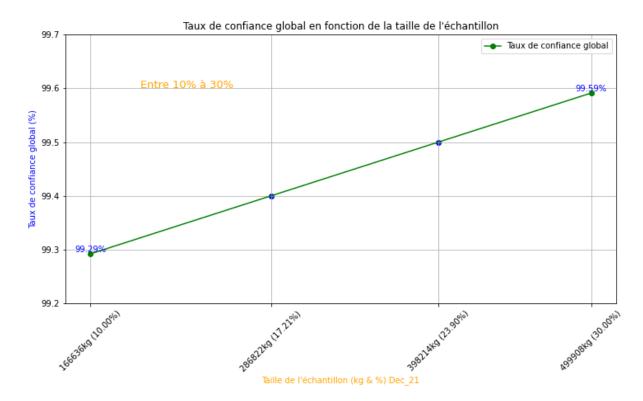


FIGURE 27 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Entre 30% à 100%:

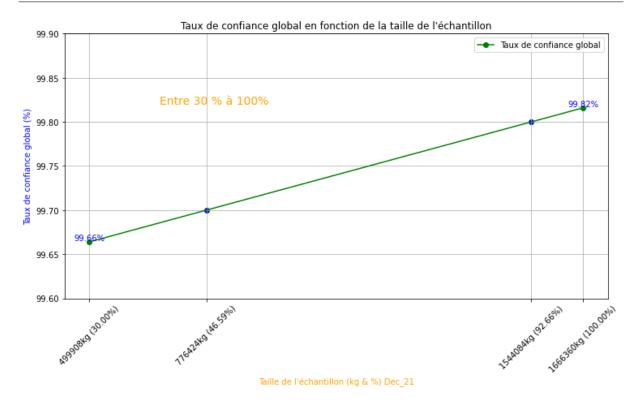


FIGURE 28 – Taux de confiance Globale en fonction de la taille de l'échantillon.

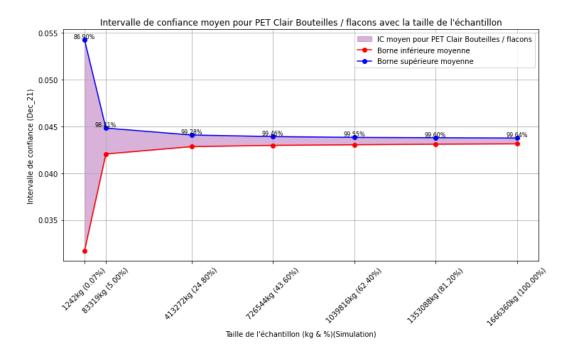


Figure 29

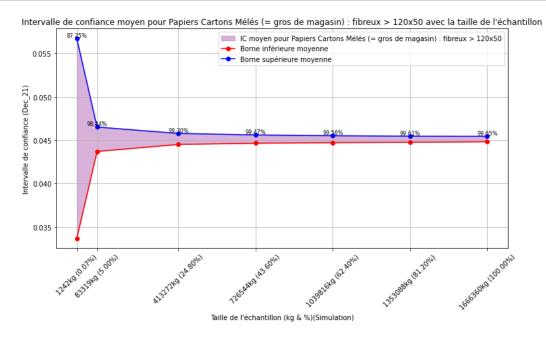


FIGURE 30

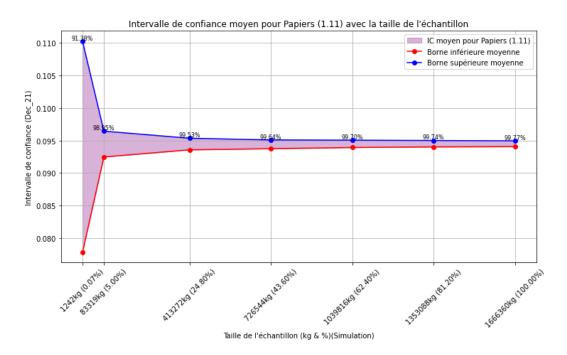


FIGURE 31

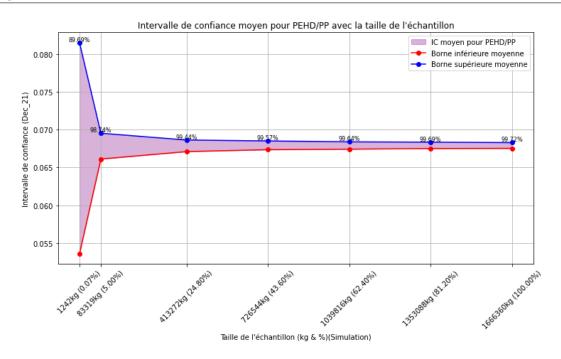


FIGURE 32

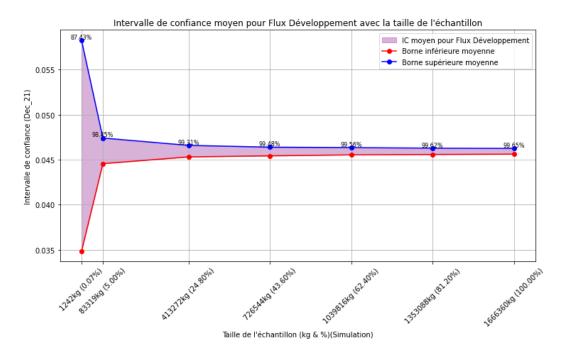


FIGURE 33

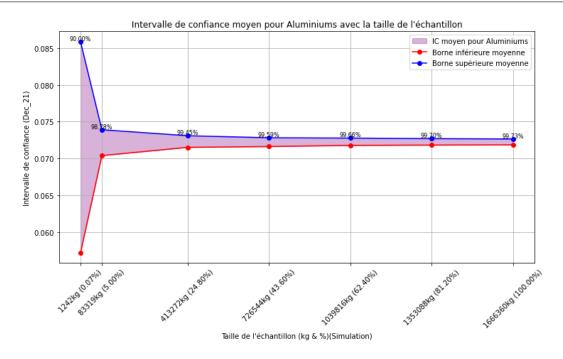


Figure 34

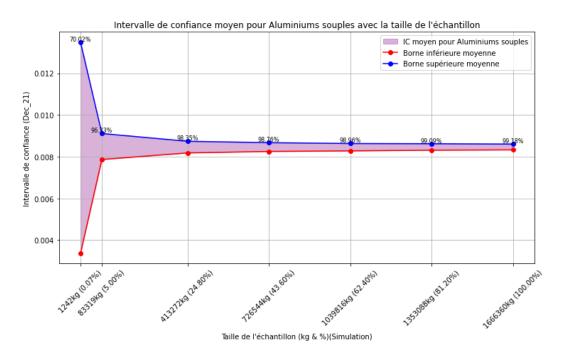


Figure 35

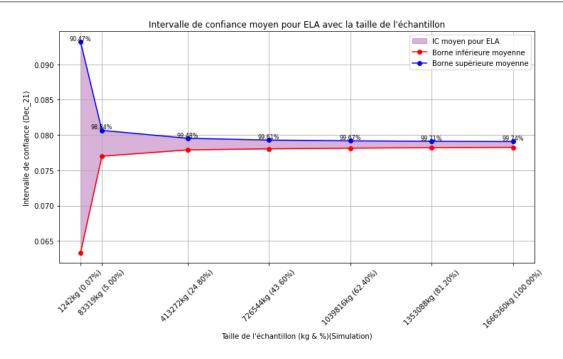


Figure 36

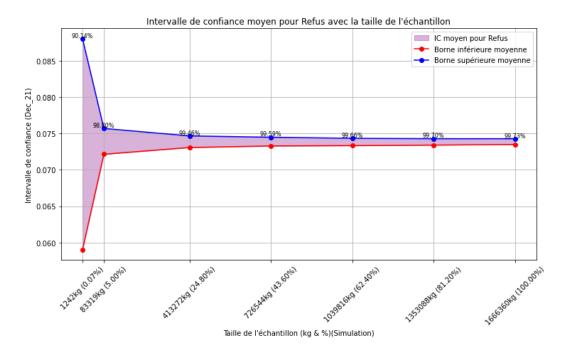


FIGURE 37

En analysant ces graphiques, on peut observer qu'au départ, avec un échantillon de 0,07% de la masse totale des déchets, le taux de confiance global commence à un peu plus de 90%. Ce taux augmente rapidement à mesure que la taille de l'échantillon s'accroît, atteignant et se maintenant à 100% pour des tailles d'échantillon plus conséquentes. Cela

indique que pour obtenir des estimations très fiables du taux de tri correct, une certaine masse critique d'échantillon est requise, au-delà de laquelle des augmentations supplémentaires de la taille de l'échantillon n'apportent pas d'amélioration significative du taux de confiance. Cela suggère qu'il est crucial de choisir une taille d'échantillon suffisante pour garantir des estimations fiables lors de l'évaluation de la qualité du tri.

VII.7 Les graphes pour Janvier 2002

Entre 10 % à 30% de la taille de l'échantillon :

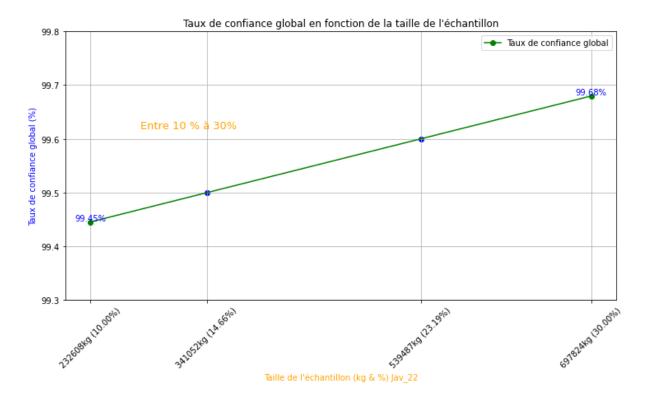


FIGURE 38 – Taux de confiance Globale en fonction de la taille de l'échantillon.

De 10% à 30% d'échantillon, on observe une augmentation de taux la confiance de 99.43% à presque 99.67%, reflétant la forte assurance obtenue avec des échantillons légèrement plus grands.

Entre 30 % à 100% de la taille de l'échantillon :

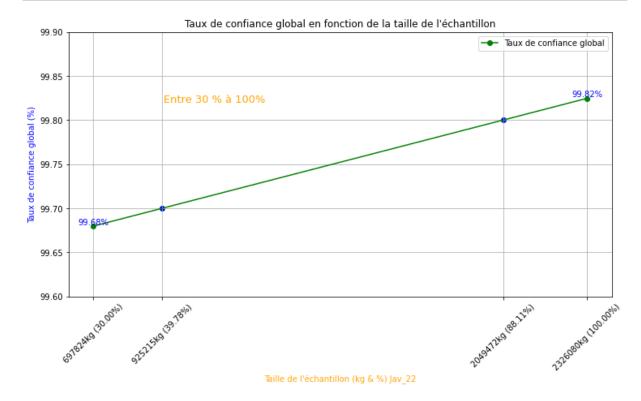


FIGURE 39 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Ce graphique montre également une augmentation constante du taux de confiance global, de près de 99.6% à pratiquement 99.8%, pour des tailles d'échantillon similaires.

La précision de l'évaluation du taux de tri s'améliore significativement avec des échantillons plus grands, indiquant que pour des évaluations fiables, il est préférable d'opter pour des tailles d'échantillon plus conséquentes. Ces résultats soulignent l'importance de choisir judicieusement la taille de l'échantillon pour garantir la validité des analyses statistiques dans des études de tri des déchets.

Les graphiques illustrent que pour obtenir un niveau de confiance global de 98% dans l'évaluation du tri des déchets pour le mois de janvier, une certaine taille minimale d'échantillon est nécessaire. On peut observer que ce niveau de confiance est largement depasser avec un échantillon de 5.00% de la masse totale des déchets, similaire à l'observation pour les autres mois.

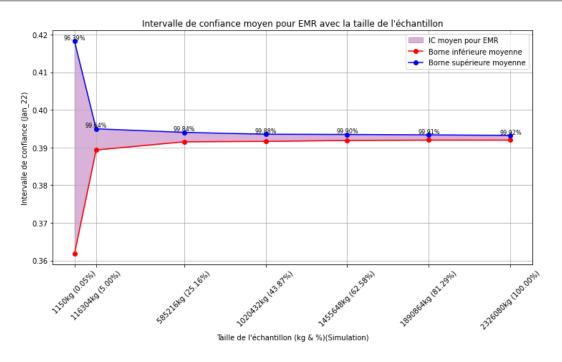


Figure 40

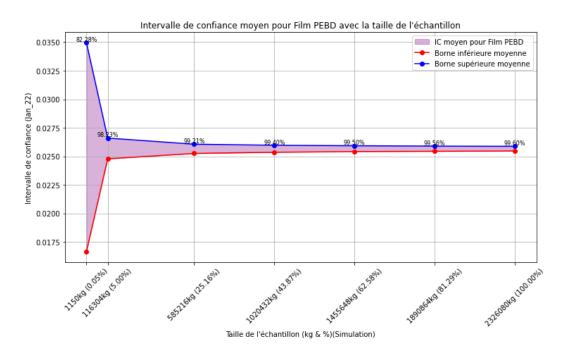


FIGURE 41

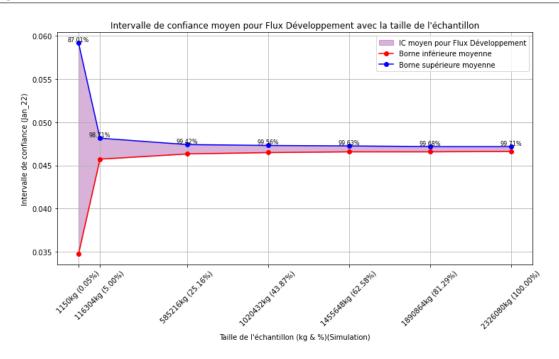


Figure 42

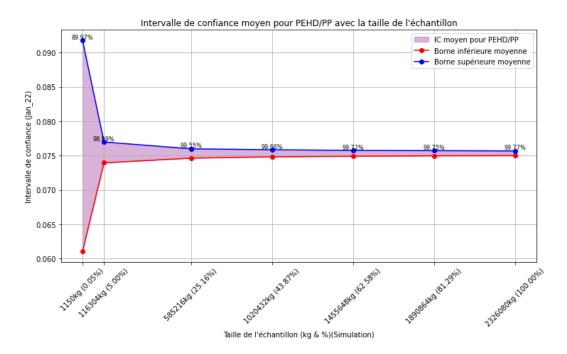


Figure 43

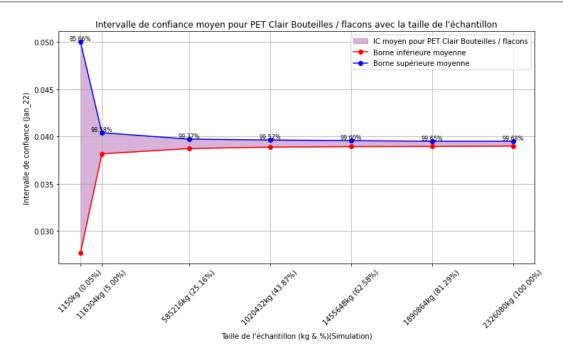


FIGURE 44

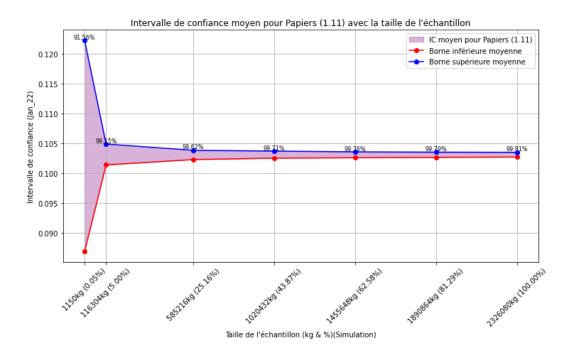


Figure 45

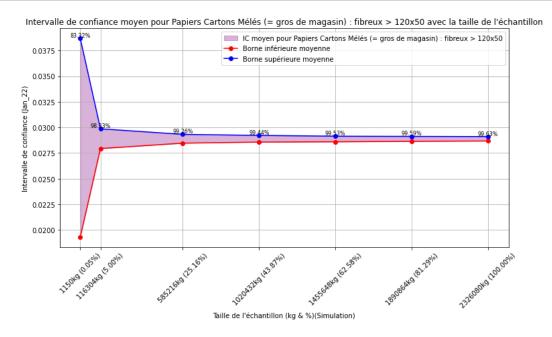


Figure 46

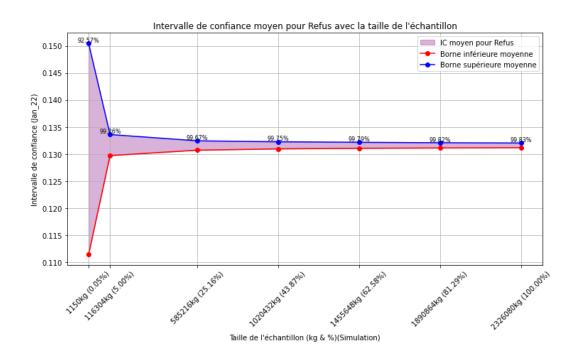


Figure 47

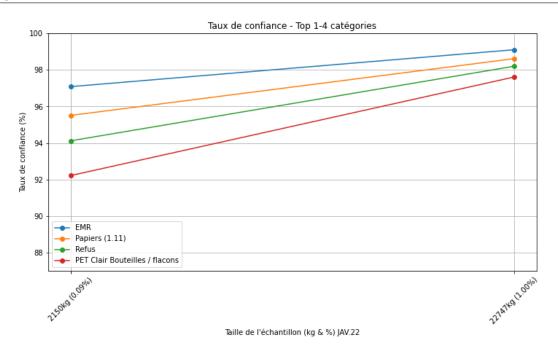


Figure 48

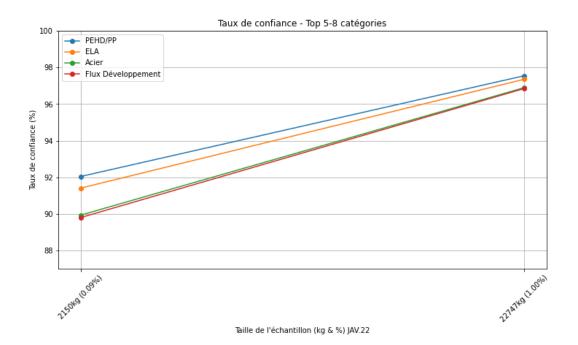


Figure 49

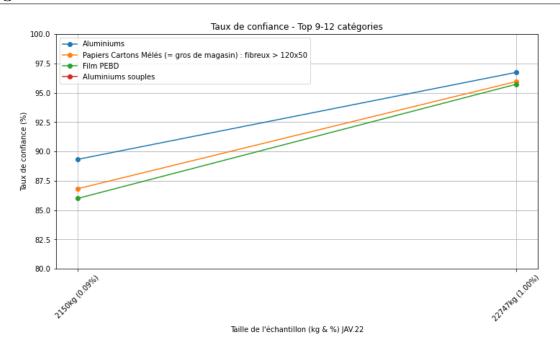


Figure 50

VII.8 Les graphes pour Février

Entre 10 % à 30% de la taille de l'échantillon :

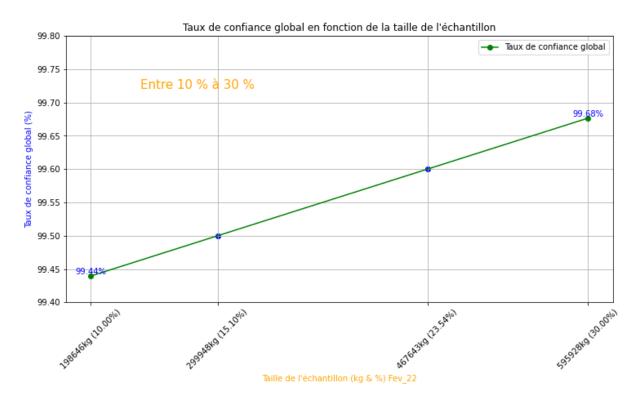


FIGURE 51 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Entre 30 % à 100% de la taille de l'échantillon :

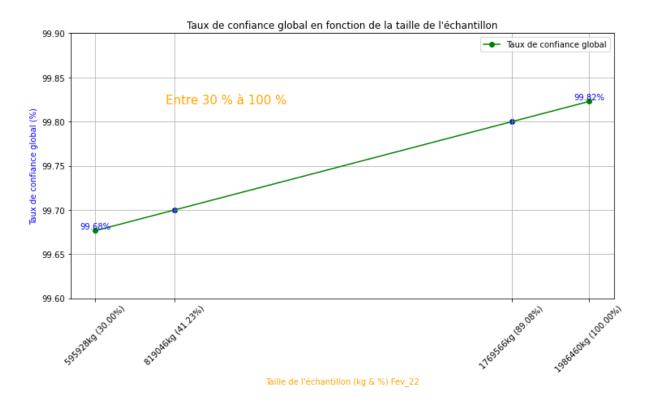


FIGURE 52 – Taux de confiance Globale en fonction de la taille de l'échantillon.

Ces analyses indiquent que des augmentations de la taille de l'échantillon améliorent systématiquement la confiance dans les estimations du tri des déchets, avec des gains de confiance plus prononcés observés aux niveaux inférieurs de la taille de l'échantillon. Pour des tailles d'échantillon très grandes, l'augmentation du taux de confiance devient moins marquée, suggérant une approche de la limite supérieure de la précision possible avec les données disponibles.

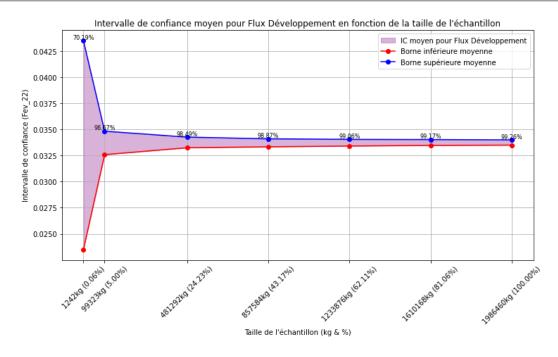


FIGURE 53

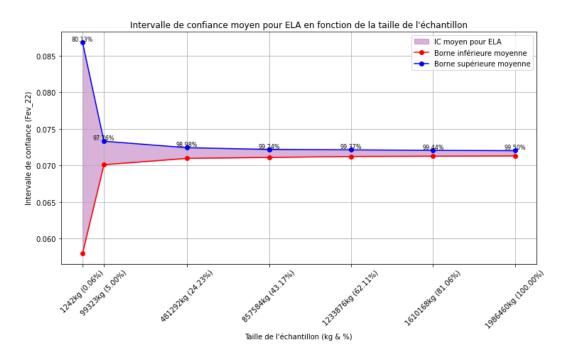


Figure 54

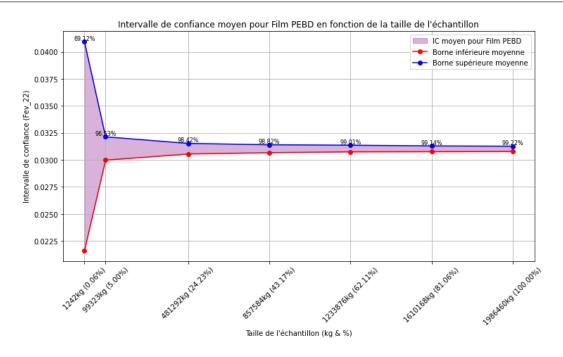


FIGURE 55

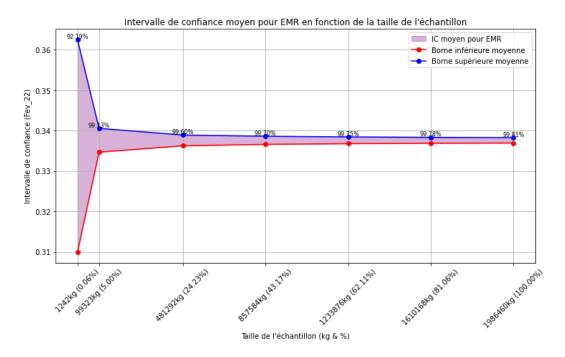


Figure 56

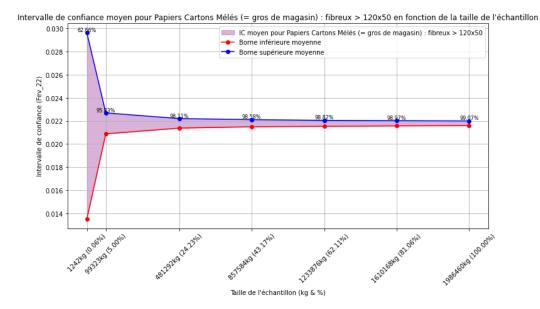


FIGURE 57

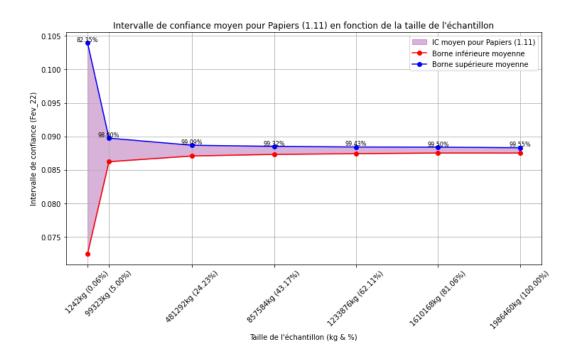


Figure 58

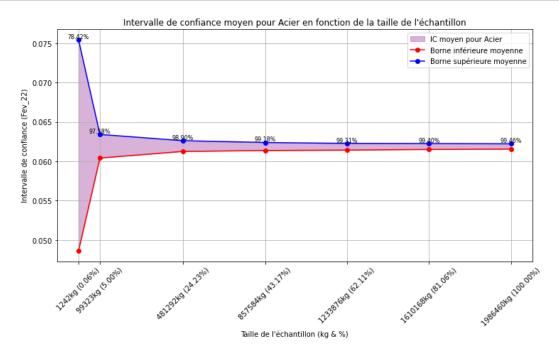


Figure 59

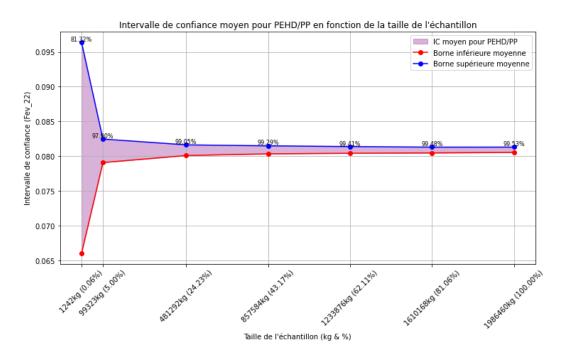


FIGURE 60

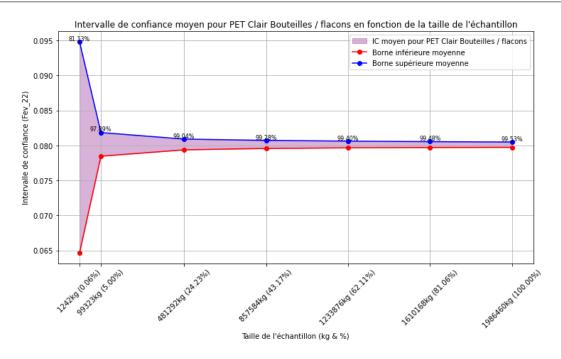


FIGURE 61

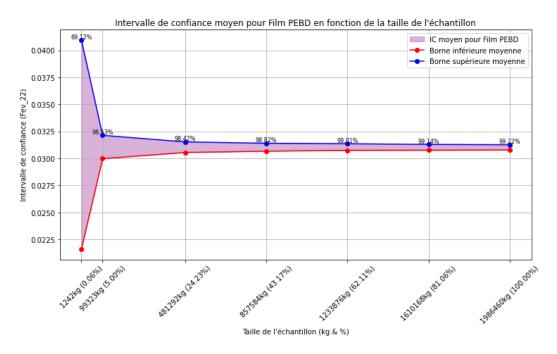


FIGURE 62

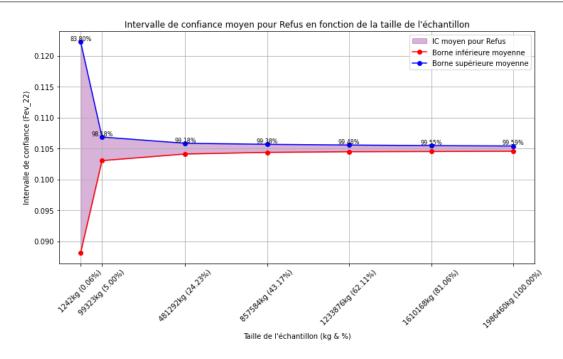


FIGURE 63

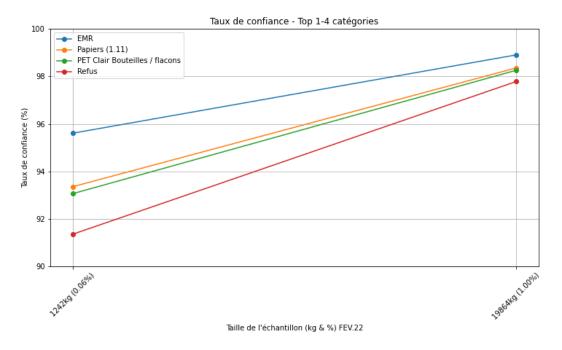


Figure 64

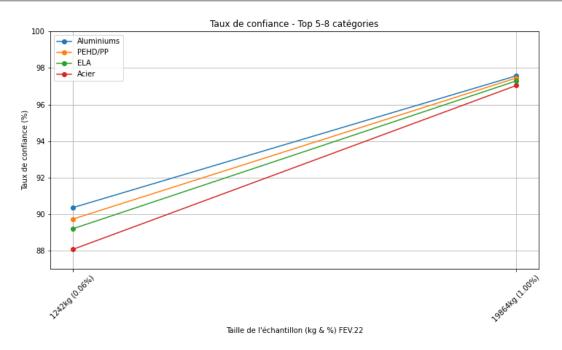


FIGURE 65

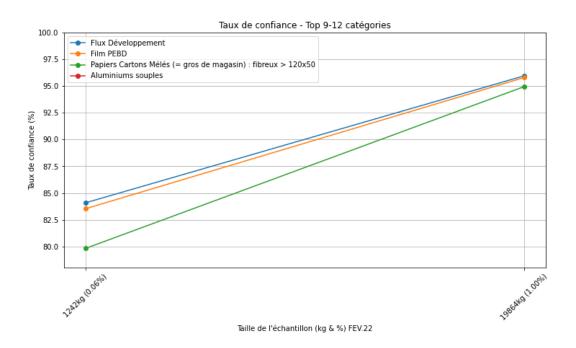


Figure 66