257

Rapports Internes de la Direction des Ressources Vivantes de l'IFREMER

ECHANTILLONNAGE DES DEBARQUEMENTS EN CRIEE

ESTIMATEURS DES CAPTURES ET VARIANCES ASSOCIEES

Dominique Pelletier

RIDRV-90.21-RH/NANTES

IFREMER

INSTITUT FRANCAIS DE RECHERCHE POUR L'EXPLOITATION DE LA MER

Adresse : IFREMER
Laboratoire ERHAL
Centre de Nantes
Rue de l'île d'Yeu - B.P. 1049
44037 Nantes Cedex 01

DIRECTION DES RESSOURCES VIVANTES

DEPARTEMENT Ressources Halieutiques
Laboratoire "Evaluation des ressources halieutiques"

STATION/LABORATOIRE Nantes

AUTEURS (S) :

Dominique PELLETIER

CODE :

DRV/RH/ERHAL
Nantes - 90/

TITRE :

ECHANTILLONNAGE DES DEBARQUEMENTS EN CRIEE, ESTIMATEURS DES CAPTURES ET VARIANCES ASSOCIEES

Date :
Tirage en nombre :

Nb pages : 30
Nb figures : 2
Nb photos :

CONTRAT
(intitulé)

N° ______

DIFFUSION
libre ☒
restreinte ☐
confidentielle ☐

RESUME

Les différents modèles d'évaluation des stocks exploités sont généralement basés sur l'analyse de captures, en âges ou en tailles, estimées par des échantillonnages dans les criées. La connaissance de la variance associée à ces estimateurs permet d'étudier l'incertitude ainsi induite sur les résultats des modèles et par conséquent sur les décisions de gestion des stocks. De plus, l'analyse des composantes de cette variance peut aider à une meilleure planification de l'échantillonnage. L'étude montre que la variance est essentiellement le fait de la disparité des débarquements selon les jours. Par ailleurs, une partie de l'effort d'échantillonnage pourrait être reportée de la catégorie commerciale correspondant à des groupes d'âge élevés, auxquels les modèles sont peu sensibles, vers les catégories économiquement plus importantes.

ABSTRACT

Stock assessment models generally rely upon catch analysis. Catches are estimated from the sampling of landings. The evaluation of the variance associated to these estimators allows to study the induced uncertainty on models results and hence on stock management decisions. Moreover, the sampling scheme may be improved by analysing the variance components. The results show that the variance is mainly due to the variability of landings between days. In addition, as assessment models show little sensitivity to old age groups, sampling effort could be partially transferred from the corresponding market category to those of major economic interest.

mots clés : Echantillonnage complexe - estimation des captures - variance.

key words : Complex sampling scheme - catch estimation - variance.

Introduction :

Les différents modèles qui décrivent la dynamique des populations exploitées par les pêcheurs nécessitent un nombre important de données. Par exemple, l'analyse de cohortes requiert la connaissance des captures selon les âges ou au moins selon les longueurs. Des échantillonnages pratiqués sur les criées permettent d'accéder à des estimations de ces données à partir de protocoles assez complexes. Il semble important d'accompagner ces valeurs des variances associées à l'échantillonnage pour avoir une idée de leur précision. Nous serons alors en mesure d'étudier les conséquences de la propagation des erreurs lors de l'utilisation des estimations dans les modèles.

Après avoir présenté un protocole en particulier, nous essaierons de prendre en compte les phases successives de l'échantillonnage afin de calculer de manière analytique des estimateurs des captures en longueur et en âge ainsi que leurs variances. On ne considérera que le calcul des variances sans s'attacher à celui de l'optimisation du plan d'échantillonnage. Ce problème a déjà fait et continue à faire l'objet de recherches particulières [1] [2]. Puis, nous appliquerons les résultats au cas de la morue pêchée en mer celtique.

1. PRESENTATION DU PROBLEME :

Nous nous intéressons aux échantillonnages réalisés par le laboratoire IFREMER de Lorient. Ces opérations sont pratiquées en routine depuis plusieurs années et selon une procédure bien précise. Les données collectées permettent d'estimer les captures en nombre, en longueur et en âge, ainsi que des paramètres caractéristiques des stocks de poisson, comme les coefficients des courbes de croissance. Les phénomènes biologiques et météorologiques induisent des effets saisonniers qualitatifs et quantitatifs importants sur les quantités débarquées. La base temporelle retenue pour les données est le **trimestre** car c'est la résolution minimale souhaitable pour les modèles.

Pratiquement, les enquêtes couvrent les criées du Sud-Ouest de la Bretagne, soit d'Audierne à Lorient et en réalité, l'essentiel concerne Douarnenez, le Guilvinec, Concarneau et Lorient qui sont visités plusieurs fois par trimestre.

Les quantités débarquées chaque jour dans ces criées sont extrêmement variables du fait de la diversité des stratégies de pêche des navires qui y sont rattachés. On en distingue en gros trois types :
- les unités côtières qui sortent pour une durée de un à quatre jours et débarquent chaque jour des quantités assez faibles et concernent principalement certaines espèces à haute valeur marchande (sole, crabes, langoustine...).
-la pêche industrielle, qui recherche principalement des espèces de moindre valeur marchande et en grande quantité, comme le lieu noir vers le Nord de l'Ecosse. Ces prises sont débarquées à part.
-les bateaux intermédiaires, qu'ils soient artisanaux ou semi-industriels qui partent pour une marée d'une durée maximale de deux semaines. Ils rentrent souvent au port le lundi et le mercredi [1] qui correspondent donc à des quantités débarquées supérieures à ceux des autres jours.

---

1 pour des raisons de distribution vers les circuits commerciaux.

Nous étudierons les débarquements de ces dernières unités car ils déterminent l'évolution de plusieurs stocks placés sous quota et concernent des secteurs géographiques cruciaux pour les pêcheries françaises. En Bretagne Sud, ces bateaux partent surtout pêcher dans le golfe de Gascogne ou en mer celtique, pour laquelle les prises françaises représentent d'ailleurs 80% des captures totales. Pour la présente étude, nous nous intéressons à cette pêcherie. Dans la mesure où la provenance du poisson débarqué est connue, cela ne pose pas de problème d'identification.

La mer celtique constitue un ensemble homogène du point de vue de la dynamique des populations de poissons. Pour plus de précisions, on se reportera à [5]. Elle s'étend au sud de l'Irlande, à l'ouest de la Grande Bretagne et de la Bretagne et est limitée à l'Ouest par l'isobathe de 600 mètres.
Les espèces pêchées sont nombreuses, on y remarque notamment des poissons démersaux comme la morue, le merlan, le merlu, et benthiques comme la baudroie, la langoustine, la raie, la cardine.

2. DESCRIPTION DE LA PROCEDURE D'ECHANTILLONNAGE :

Nous allons simultanément présenter le protocole et justifier les hypothèses de modélisation de l'échantillonnage qui en découlent. Les rappels statistiques nécessaires sont présentés en annexe.

2.1. Choix des jours :

Avant chaque trimestre, un certain nombre de jours d'échantillonnage sont décidés pour chaque catégorie commerciale de chaque espèce. Par la suite, ces jours sont choisis parmi les jours de débarquement des unités. Ce choix dépend probablement de la disponibilité des enquêteurs et d'autres facteurs non apparents et il est aussi soumis à des problèmes de coût. Ainsi, puisque la criée affiche à l'avance les bateaux qui rentrent, il est possible d'avoir une idée des quantités qui vont être débarquées. Néanmoins, puisque les influences de ces facteurs ne sont pas démontrées ni quantifiables, nous considérerons que, pour chaque trimestre, les jours d'enquête sont tirés **au hasard** parmi les jours de débarquements dans la criée concernée. Ces jours ainsi que les quantités correspondantes sont recensés dans une base de données statistiques.

2.2. Les catégories commerciales et la répartition en caisses :

Considérons maintenant la manipulation effectuée un jour donné. La procédure ne prend pas en compte un niveau bateau car cela nécessiterait un effort d'échantillonnage trop important. Nous considérerons donc les débarquements tous bateaux confondus. A la criée, les poissons sont classés par espèces, puis pour chaque espèce, par catégories commerciales en nombre variable d'une espèce à l'autre, mais constant pour une espèce donnée. La définition de ces catégories est dictée par des normes européennes, basées sur le poids du poisson (cf. tableau 1).

Catégorie commerciale	1	2	3	4	5
Kg/individu	7 et +	4 à 7 exclu	2 à 4 exclu	1 à 2 exclu	0.3 à 1 exclu

Tableau 1. Définition des catégories commerciales européennes pour la morue par le règlement R(CEE) 103/76 modifié R(CEE) 3940/87.

Au sein de chaque catégorie commerciale, les poissons sont toujours rangés **par caisses** de poids prédéfini, par exemple, 50 kg à Lorient, 20 ou 50 kg selon l'espèce au Guilvinec.

En effet, encore actuellement, la plupart des criées mettent en vente des *lots de caisses* et non des poids de poisson. Seuls les ports de Lorient et de St Guénolé effectuent des pesées systématiques. La criée de ce dernier est en effet très récente et la prépondérance dans les débarquements des langoustines, à forte valeur marchande, explique ces installations particulières.

La répartition des poissons en catégories commerciales est pratiquée empiriquement par les professionnels (pêcheurs et/ou personnel de la criée). Ce classement est fondé sur la taille et le poids de l'animal. Les poissons blessés ou d'aspect non commercialisable sont déclassés en deuxième ou troisième choix. Ils représentent une part très faible du total et nous les négligerons devant les autres, ils ne sont d'ailleurs pas échantillonnés. A cause de la nature subjective de la répartition, chaque catégorie correspond à une gamme assez large de tailles et si l'on compare *a posteriori* les histogrammes en longueurs des catégories, on constate un fort recouvrement.

Ainsi, un poisson d'une taille donnée est affecté à une catégorie, mais s'il avait été débarqué un autre jour, il se serait peut-être retrouvé dans une catégorie adjacente (voire dans une autre plus éloignée), soit à cause d'une personne qui trie différemment, soit parce que le tri est influencé par l'abondance des débarquements. Au niveau trimestriel, les catégories commerciales ne sont donc pas définies *a priori* sur l'ensemble des poissons débarqués. Par contre, au plan journalier, chaque individu débarqué appartient à **une classe et une seule.** Cependant, dans la mesure où chaque catégorie commerciale n'est pas échantillonnée le même nombre de jours qu'une autre dans le trimestre, nous devrons supposer que ce sont les jours de débarquements d'une espèce donnée qui sont **stratifiés** par catégorie commerciale (et non les caisses d'un jour donné). La taille de chacune des strates est **connue** puisque le nombre de caisses ou la quantité débarquées dans chaque catégorie sont enregistrés à la criée pour chaque espèce et chaque jour. Cette stratification des captures est susceptible d'apporter un gain de précision à l'estimation des effectifs en longueur et en âge du fait de la forte corrélation existant entre l'appartenance à une catégorie commerciale et la taille ou l'âge d'un poisson.
Quant aux caisses, elles sont remplies durant le tri en catégories commerciales. Il est raisonnable d'admettre qu'elles sont formées **au hasard** à l'intérieur d'une catégorie commerciale, c'est à dire que les caractéristiques moyennes en nombre, en taille, et en âge des poissons de chaque caisse d'une espèce donnée seront identiques pour un jour et une catégorie donnés.

2.3. L'échantillonnage des caisses :

Dans chaque catégorie commerciale, on échantillonne un certain nombre de caisses prises **au hasard**. Dès lors, la manipulation comporte trois volets. D'abord, on dénombre exhaustivement le nombre d'individus contenus dans chaque caisse tirée. La valeur obtenue est donc **déterminée** si l'on suppose qu'il n'y a pas d'erreur de comptage. Simultanément, on mesure un poisson **systématiquement** tous les un ou deux [2] poissons comptés. Enfin, pour chaque longueur, exprimée en centimètres, on prélève, également de manière **systématique,** un otolithe (ou un rayon de nageoire selon les espèces) tous les un, deux ou quatre individus mesurés pour une longueur donnée. Ces otolithes serviront à déterminer l'âge de l'animal. La lecture de l'âge des individus d'une espèce donnée entraîne à la longue des erreurs systématiques et corrélées –une lecture d'âge est toujours plus ou moins influencée par les précédentes–. Des études sont actuellement menées à ce sujet. De plus, un certain nombre d'otolithes sont illisibles ou ratés lors de la préparation, il est toutefois possible d'en connaître la quantité et d'obtenir finalement le **taux réel** d'âges déterminés. Ainsi, nous admettrons que **tous** les otolithes prélevés sont lus **sans erreur.** Par ailleurs, puisqu'une **proportion donnée** d'otolithes est prélevée pour **chaque** longueur échantillonnée, la longueur d'un poisson dont on détermine l'âge n'est connue qu'après le prélèvement. Pour la détermination des effectifs aux âges, les poissons d'une caisse peuvent être **stratifiés a posteriori** selon les longueurs.

Ces trois séries de données permettent de produire les estimations des effectifs débarqués pendant le trimestre dans cette criée, pour une longueur et pour un âge donnés. Par sommation, nous pouvons ensuite en déduire les effectifs correspondants tous ports confondus.

3. MODELISATION DE L'ECHANTILLONNAGE :

Des modélisations simplifiées existent mais elle ne prennent pas en compte toutes les phases de l'échantillonnage et n'ont semble-t-il pas fait l'objet d'une publication. Le protocole précédemment décrit met en évidence plusieurs **niveaux** d'échantillonnage et une **stratification.** Les notations nécessaires sont résumées ci-dessous.

Notations :

Les indices sont définis comme suit :

j pour le jour échantillonné.
h pour la catégorie commerciale correspondante.
i pour la caisse.
l pour la longueur.
a pour l'âge.

2 Ce taux varie uniquement selon la catégorie commerciale. Par exemple, dans la catégorie 1, les gammes de taille et d'âge représentées sont beaucoup plus larges que dans les autres catégories. A précision égale, une estimation des effectifs pour chaque longueur et chaque âge dans une catégorie donnée requiert donc un échantillon plus grand.

	Population	Echantillon
Nombre total de jours de débarquement	N	n
Jours de débarquement dans la strate h	N_h	n_h
Nombre de caisses du jour j dans la strate h	M_{hj}	m_{hj}

En premier lieu, les catégories commerciales forment des strates à l'intérieur du trimestre. Considérons ensuite ce que nous appellerons le niveau "jours" : chaque strate se compose d'**unités primaires** : les jours de débarquement. Chaque jour est lui-même constitué d'**unités secondaires** : les caisses débarquées ce jour dans la catégorie commerciale considérée. Ces caisses pèsent en principe 50 kg et le nombre de caisses ou la quantité débarquées chaque jour, très variables, sont enregistrés à la criée. Les unités primaires sont donc de **tailles connues et fortement inégales** à l'intérieur d'une strate et les unités secondaires sont de **tailles égales.**

Par conséquent, lorsque nous estimons le nombre total de poissons débarqués pendant le trimestre, sans distinction de longueur ni d'âge, l'échantillonnage peut être modélisé comme suit :

- Echantillonnage aléatoire stratifié de n_h unités primaires sur un total de N_h jours de débarquement dans la catégorie h du trimestre.
- Echantillonnage aléatoire simple (EAS) avec un taux m_{hj}/M_{hj} pour chaque couple jour-strate sur les caisses.

Chaque caisse est ici l'unité statistique sur laquelle la variable mesurée est le nombre total de poissons qu'elle contient.

Considérons maintenant la phase suivante : Parmi les poissons dénombrés, une certaine proportion f est mesurée systématiquement et pour un certain nombre de poissons d'une longueur donnée, un otolithe est prélevé. Les deux proportions dépendent de la catégorie commerciale. Ces échantillonnages systématiques produisent des variances inférieures ou égales à celle de l'échantillonnage aléatoire simple (EAS) [4]. Cependant, cette variance n'est pas calculable simplement. Par contre, nous pouvons soit négliger la variabilité intra-caisse en supposant qu'on trouve les valeurs réelles en multipliant par deux ce que l'on compte, soit considérer qu'on a des EAS de taux d'échantillonnage respectifs f_h parmi les poissons d'une caisse et f'_h parmi les poissons de longueur l d'une caisse pour une catégorie h donnée.[3] Dans un premier temps, nous nous placerons sous cette deuxième hypothèse, c'est à dire qu'au pire, nous surestimerons la variance.

Par ailleurs, dans la mesure où l'on prélève un otolithe pour un nombre fixé d'animaux de longueur l mesurés, nous pouvons prendre en compte une **poststratification** [4] selon les longueurs à l'intérieur des caisses. Celle-ci suppose la connaissance **a priori** des tailles des strates, ce qui n'est pas le cas. L'estimation des tailles des strates peut aussi faire l'objet d'un double échantillonnage [4]. Afin de ne pas alourdir les calculs,

3 Les taux d'échantillonnage peuvent varier. Ils seront notés f et f' pour garder un caractère général au calcul.

nous ne tiendrons pas compte de cette possibilité et donc nous ne tirerons pas parti de la relation entre l'âge et la taille. Si nous ignorons cette relation, nous pouvons par conséquent considérer que l'échantillonnage des otolithes est un échantillonnage **aléatoire simple de taux** $f_h f'_h$ à l'intérieur d'une caisse. Les estimations des effectifs aux âges et aux longueurs sont corrélées, mais nous ne connaîtrons pas leurs covariances. Cependant, dans la mesure où nous n'utilisons pas les estimations en longueur et en âge **simultanément** dans nos modèles, cela n'aura pas de conséquences incontrôlées. Par contre, pour ajuster une courbe de croissance et une relation taille-poids ou établir une clé taille-âge, il serait souhaitable d'en tenir compte.

Finalement, l'estimation du nombre de poissons de longueur l du trimestre résultera du modèle suivant :

- Echantillonnage aléatoire stratifié de n_h unités primaires sur un total de N_h jours de débarquement dans la strate h du trimestre.
- EAS avec un taux m_{hj}/M_{hj} pour chaque couple jour-strate sur les unités secondaires que sont les caisses.
- EAS au taux f_h de poissons mesurés parmi les poissons comptés.

Et pour l'estimation du nombre de poissons d'âge a du trimestre, on aura un modèle très peu différent :

- Echantillonnage aléatoire stratifié de n_h unités primaires sur un total de N_h jours de débarquement dans la strate h du trimestre.
- EAS avec un taux m_{hj}/M_{hj} pour chaque couple jour-strate sur les unités secondaires que sont les caisses.
- EAS au taux $f_h f'_h$ de poissons âgés parmi les poissons comptés.

En dernier lieu, il nous faut une hypothèse supplémentaire pour pouvoir estimer le nombre d'individus de longueur l et d'âge a. Par exemple, lorsque l'on effectue l'EAS pour les longueurs, on dispose des données $D_{jhil}, j = 1..n, h = 1..H, l = 1..L, i = 1..mjh$. D_{jhil} représente le nombre d'individus de longueur l dans la caisse i de la catégorie commerciale h correspondant au jour j. Dans chaque caisse, il existe une certaine distribution des longueurs caractérisée par les proportions $p_{il}, l = 1..L$. Nous supposerons que le tirage effectué dans cette caisse est **aléatoire** avec un taux d'échantillonnage f_h. Le vecteur $(D_{jhil}, D_{jhi2}...D_{jhiL})$ suit la **loi multinomiale** [4] de paramètres $(D_{jhi}, p_{il}, p_{i2}...p_{iL})$.
En ce qui concerne les âges, le raisonnement est analogue. D_{jhia} représente le nombre d'individus d'âge a dans la caisse i de la catégorie commerciale h correspondant au jour j. La distribution des âges dans la caisse i est caractérisée par les proportions $p_{ia}, a = 1..A$. On admettra de même que le vecteur $(D_{jhil}, D_{jhi2}...D_{jhiA})$ suit la **loi multinomiale** de paramètres $(D_{jhi}, p_{il}, p_{i2}...p_{iA})$.
Cette hypothèse nous permettra d'obtenir des variances et des covariances entre les estimateurs pour deux longueurs ou deux âges différents.

Grâce au modèle statistique que nous venons de formaliser, nous allons maintenant exprimer les estimateurs et leurs variances, avant de procéder à une application numérique dans le cas des débarquements de morue à Lorient.

4 Du fait du tirage sans remise, c'est en fait une loi multihypergéométrique, mais sa variance diffère seulement de celle de la loi multinomiale par la correction de population finie.[4]

4. EXPRESSIONS DES ESTIMATEURS :

4.1. Les estimateurs finaux :

Nous cherchons à estimer :

N_{tot} le nombre total de poissons du trimestre.

N_l le nombre total de poissons de longueur l du trimestre.

N_a le nombre total de poissons d'âge a du trimestre.

Nous commencerons le calcul des estimateurs de la moyenne et de la variance de chacune de ces quantités par le niveau le plus externe de l'échantillonnage, c'est à dire les jours. Par ailleurs, lorsqu'elles ne sont pas explicitées, les expressions des covariances sont strictement analogues à celles des variances. En ce qui concerne les notations, nous utiliserons la variable X pour toutes les quantités intermédiaires. L'indice + représente une sommation sur tous les indices non représentés, son absence signifie qu'il s'agit d'un nombre moyen. Par exemple :

X_{h+} est le nombre total d'animaux de la strate h.

X_{hj} est le nombre moyen par caisse d'animaux du couple jour–strate (j,h).

Par conséquent, on a les égalités :

$$N_{tot} = X_+$$

$$N_l = X_{l+}$$

$$N_a = X_{a+}$$

4.2. Le niveau jour :

Il s'agit d'un échantillonnage stratifié de n_h jours sur un total de N_h jours par strate du trimestre. Les estimateurs finaux recherchés s'écrivent donc :

$$\hat{X}_+ = N\hat{X} = \sum_{h=1}^{H} N_h X_h$$

$$\hat{X}_{l+} = \sum_{h=1}^{H} N_h X_{lh}$$

$$\hat{X}_{a+} = \sum_{h=1}^{H} N_h X_{ah}$$

Leurs variances s'exprimeront par :

$$Var(\hat{X}_{+}) = \sum_{h=1}^{H} N_h^2 Var(X_h)$$

$$Var(\hat{X}_{l+}) = \sum_{h=1}^{H} N_h^2 Var(X_{lh})$$

$$Var(\hat{X}_{a+}) = \sum_{h=1}^{H} N_h^2 Var(X_{ah})$$

Les différentes moyennes par catégorie sont respectivement estimées par :

$$\hat{X}_h = \frac{1}{n_h}\sum_{j=1}^{n_h} X_{hj+} \qquad \hat{X}_{lh} = \frac{1}{n_h}\sum_{j=1}^{n_h} X_{lhj+} \qquad et \qquad \hat{X}_{ah} = \frac{1}{n_h}\sum_{j=1}^{n_h} X_{ahj+}$$

Les variances de ces estimateurs s'obtiennent en considérant les deux premiers niveaux d'échantillonnage. Si nous nous conformons à la formule explicitée en annexe, §III. , la variance de l'estimateur du nombre de poissons de la strate h s'exprime par :

$$Var(\hat{X}_h) = Var_{jours}\{E_{caisses}(\hat{X}_h)\} + E_{jours}\{Var_{caisses}(\hat{X}_h)\}$$

En se reportant à l'annexe, §III., on montre que cette variance peut être estimée par :

$$V\hat{a}r(\hat{X}_h) = \frac{N_h - n_h}{N_h n_h (n_h - 1)}\sum_{j}^{n_h}(\hat{X}_{hj} - \hat{X}_h)^2 + \frac{1}{n_h N_h}\sum_{j}^{n_h} M_{hj}^2 Var_{caisses}(\hat{X}_{hj})$$

Soit :

$$V\hat{a}r(\hat{X}_{+}) = \sum_{h=1}^{H}\left[\frac{(N_h - n_h)N_h}{n_h(n_h-1)}\sum_{j}^{n_h}(\hat{X}_{hj} - \hat{X}_h)^2 + \left(\frac{N_h}{n_h}\right)\sum_{j}^{n_h} M_{hj}^2 Var_{caisses}(\hat{X}_{hj})\right]$$

Les expressions sont tout à fait analogues pour $\hat{X}_{l+}$, et $\hat{X}_{a+}$. Les termes de covariance entre deux longueurs ou deux âges différents s'obtiennent aussi de la même manière :

–Pour les estimateurs finaux :

$$Cov(\hat{X}_{l1+}, \hat{X}_{l2+}) = \sum_{h=1}^{H} N_h^2 Cov(\hat{X}_{l1h}, \hat{X}_{l2h})$$

–Pour les estimateurs des moyennes par strate :

$$Cov(\hat{X}_{l1h}, \hat{X}_{l2h}) = Cov_{jours}\{E_{caisses}(\hat{X}_{l1h}), E_{caisses}(\hat{X}_{l2h})\} + E_{jours}\{Cov_{caisses}(\hat{X}_{l1h}, \hat{X}_{l2h})\}$$

Ce terme de covariance s'estime comme celui de la variance (voir annexe, §III.).

Nous devons maintenant calculer les variances et covariances à l'intérieur d'un jour et par là même nous intéresser à l'échantillonnage des caisses.

4.3. L'échantillonnage des caisses :

Considérons d'abord l'estimation du nombre total de poissons du trimestre. Pour chaque catégorie commerciale h et chaque jour j, un EAS de m_{hj} caisses sur un total de M_{hj}, permet d'estimer X_{hj}, le nombre moyen de poissons pour le couple jour–strate par :

$$\hat{X}_{hj} = \frac{\sum_{i}^{m_{hj}} X_{hji}}{m_{hj}}$$

où X_{hji} est le nombre de poissons comptés dans la caisse i, soit son effectif total. Nous avons supposé cette valeur connue sans erreur, c'est pourquoi le X ne porte pas de chapeau.

La variance de cet estimateur est celle d'un EAS (voir annexe, §I.), estimée par :

$$V\hat{a}r(\hat{X}_{hj}) = \left(\frac{M_{hj} - m_{hj}}{M_{hj}}\right)\frac{s_{hj}^2}{m_{hj}}$$

où s_{hj}^2 vaut :

$$s_{hj}^2 = \frac{\sum_{i=1}^{m_{hj}} (X_{hji} - \hat{X}_{hj})^2}{m_{hj} - 1} .$$

A partir de cette variance, on peut en déduire celle de $\hat{X}_h$, et finalement celle de $\hat{X}_+$. Ainsi, le nombre total de poissons du trimestre est estimé par la quantité :

$$\hat{X}_+ = \sum_{h}^{H} \frac{N_h}{n_h}\left(\sum_{j}^{n_h} \frac{M_{hj}}{m_{hj}} \sum_{i}^{m_{hj}} X_{hji}\right)$$

La variance de $\hat{X}_+$ est finalement estimée par :

$$\boxed{\begin{aligned} V\hat{a}r(\hat{X}_+) = & \sum_h^H \frac{N_h(N_h - n_h)}{n_h(n_h-1)} \sum_j^{n_h} (\hat{X}_{hj} - \hat{X}_h)^2 \\ & + \sum_h^H \left(\frac{N_h}{n_h}\right) \sum_j^{n_h} \frac{M_{hj}(M_{hj} - m_{hj})}{m_{hj}(m_{hj}-1)} \left(\sum_{i=1}^{m_{hj}} (\hat{X}_{hji} - \hat{X}_{hj})^2 \right) \end{aligned}}$$

En ce qui concerne les effectifs en longueur et en âge, le calcul débute de manière identique et on obtient les estimateurs des moyennes par strate et par jour :

$$\hat{X}_{lhj} = \frac{\sum_i^{m_{hj}} X_{lhji}}{m_{hj}}$$

$$\hat{X}_{ahj} = \frac{\sum_i^{m_{hj}} X_{ahji}}{m_{hj}}$$

où $\hat{X}_{lhji}$ et $\hat{X}_{ahji}$ sont les nombres de poissons respectifs de longueur l et d'âge a dans la caisse i.

Cependant, si l'estimation des effectifs globaux résulte d'un EAS sur les caisses, l'échantillonnage de ces caisses constitue un niveau supplémentaire lorsque l'on s'intéresse aux effectifs en longueur et en âge. Ceci ajoute une composante à la variance. Par le même raisonnement que pour le niveau jour, nous trouvons une estimation de la variance (se reporter à l'annexe, §III.) :

$$V\hat{a}r(\hat{X}_{lhj}) = \left(\frac{M_{hj} - m_{hj}}{M_{hj}} \frac{s^2_{lhj}}{m_{hj}} \right) + \frac{1}{m_{hj} M_{hj}} \sum_{i=1}^{m_{hj}} Var(\hat{X}_{lhji})$$

où s^2_{lhj} vaut :

$$s^2_{lhj} = \frac{\sum_{i=1}^{m_{hj}} (\hat{X}_{lhji} - \hat{X}_{lhj})^2}{m_{hj} - 1}$$

De même pour les âges, on a :

$$V\hat{a}r(\hat{X}_{ahj}) = \left(\frac{M_{hj} - m_{hj}}{M_{hj}} \frac{s^2_{ahj}}{m_{hj}} \right) + \frac{1}{m_{hj} M_{hj}} \sum_{i=1}^{m_{hj}} Var(\hat{X}_{ahji})$$

où s^2_{ahj} vaut :

$$s^2_{ahj} = \frac{\sum_{i=1}^{m_{hj}} (\hat{X}_{ahji} - \hat{X}_{ahj})^2}{m_{hj} - 1}$$

L'expression des variances à l'intérieur de chaque caisse reste à calculer. Les estimations des effectifs de longueur l et d'âge a dans une caisse proviennent d'EAS de taux différents (voir §3).

4.4. Les échantillonnages en longueur et en âge :

L'expérimentateur effectue maintenant dans chaque caisse des EAS de taux f_h pour les longueurs et $f_h f'_h$ pour les âges. Si D_{lhji} représente le nombre de poissons de longueur l tirés dans la caisse i, nous estimerons le nombre de poissons de longueur l total dans cette caisse par :

$$X_{lhji} = \frac{D_{lhji}}{f_h}$$

Nous avons également supposé que le nombre de poissons d'une longueur donnée obtenu dans une caisse par cet échantillonnage suivait une loi multinomiale. Par conséquent, si l'on se réfère à l'annexe, §IV., la variance de $\hat{X}_{jhli}$ s'écrit :

$$Var(\hat{X}_{lhji}) = \frac{(1 - f_h)}{f_h X_{hji}} p_{lhji}(1 - p_{lhji}) X^2_{hji} = \left(\frac{1 - f_h}{f_h}\right) p_{lhji}(1 - p_{lhji}) X_{hji}$$

où p_{lhji} représente la probabilité de tirer un poisson de longueur l dans la caisse i pour le jour j et la catégorie h.
De même, la covariance entre les estimateurs pour deux longueurs différentes vaut :

$$Cov(\hat{X}_{l1hji}, \hat{X}_{l2hji}) = -\left(\frac{1 - f_h}{f_h}\right) p_{l1hji} p_{l2hji} X_{hji}$$

Ces expressions seront estimées en remplaçant p_{lhji} par son estimation dans l'échantillon :

$$\hat{p}_{lhji} = \frac{\hat{X}_{lhji}}{\hat{X}_{hji}}$$

Soit après simplifications :

$$V\hat{a}r(\hat{X}_{lhji}) = \frac{(1-f_h)}{f_h}\frac{\hat{X}_{lhji}(\hat{X}_{hji}-\hat{X}_{lhji})}{\hat{X}_{hji}}$$

$$C\hat{o}v(\hat{X}_{l1hji},\hat{X}_{l2hji}) = -\left(\frac{1-f_h}{f_h}\right)\frac{\hat{X}_{l1hji}\hat{X}_{l2hji}}{\hat{X}_{hji}}$$

En remplaçant les estimateurs des variances et des covariances du §4.3. par les expressions calculées dans le développement précédent, on obtiendra la formule :

$$Var(\hat{X}_{lhj}) = \frac{M_{hj}(M_{hj}-m_{hj})}{(m_{hj}-1)m_{hj}}\sum_{i=1}^{m_{hj}}(\hat{X}_{lhji}-\hat{X}_{lhj})^2 + \left(\frac{1-f_h}{f_h}\right)\left(\frac{M_{hj}}{m_{hj}}\right)\sum_{i=1}^{m_{hj}}\frac{\hat{X}_{lhji}(\hat{X}_{hji}-\hat{X}_{lhji})}{\hat{X}_{hji}}$$

L'expression est strictement analogue pour la covariance. On obtient les formules finales en substituant cette expression dans la formule établie au §4.2. La même démarche est adoptée pour l'estimation des effectifs en âges. Ici encore, on admet que le nombre de poissons d'un âge donné obtenu par échantillonnage suit une loi multinomiale. Ainsi, conformément à l'annexe, §IV., la variance de l'estimateur du nombre total de poissons d'âge a dans la caisse i noté $\hat{X}_{ahji}$ s'écrit :

$$Var(\hat{X}_{ahji}) = (1-f_h f'_h)\frac{p_{ajhi}(1-p_{ajhi})}{f_h f'_h X_{hji}}X_{hji}^2 = \frac{(1-f_h f'_h)}{f_h f'_h}p_{ajhi}(1-p_{ajhi})X_{hji}$$

De même, le terme de covariance entre deux âges différents vaut :

$$Cov(\hat{X}_{a1hji},\hat{X}_{a2hji}) = -\left(\frac{1-f_h f'_h}{f_h f'_h}\right)p_{a1hji}p_{a2hji}X_{hji}$$

Ces deux quantités seront estimées en remplaçant p_{ajhi} par son estimation dans l'échantillon :

$$\hat{p}_{ajhi} = \frac{D_{ahji}}{D_{hji}} = \frac{\hat{X}_{ahji}}{\hat{X}_{hji}}$$

Soit :

$$V\hat{a}r(\hat{X}_{ahji}) = \left(\frac{1-f_h f'_h}{f_h f'_h}\right)\frac{\hat{X}_{ahji}(\hat{X}_{hji}-\hat{X}_{ahji})}{\hat{X}_{hji}}$$

et :

$$C\hat{o}v(\hat{X}_{a1hji}, \hat{X}_{a2hji}) = -\left(\frac{1-f_h f'_h}{f_h f'_h}\right)\frac{\hat{X}_{a1hji}\hat{X}_{a2hji}}{\hat{X}_{hji}}$$

Il reste à remplacer les variances dans les formules établies au §4.3. par leurs expressions *in extenso* pour parvenir à :

$$Var_{caisses}(\hat{X}_{ahj}) = \frac{M_{hj}(M_{hj}-m_{hj})}{m_{hj}(m_{hj}-1)}\sum_{i=1}^{m_{hj}}(\hat{X}_{ahji}-\hat{X}_{ahj})^2$$
$$+\left(\frac{1-f_h f'_h}{f_h f'_h}\right)\left(\frac{M_{hj}}{m_{hj}}\right)\sum_{i=1}^{m_{hj}}\hat{X}_{ahji}(\hat{X}_{hji}(\hat{X}_{hji}-\hat{X}_{ahji}))$$

Comme pour les longueurs, l'expression établie au §4.2. permet alors d'obtenir l'expression finale de la variance. La covariance se calcule de la même manière.

4.5. Les expressions finales des estimateurs et de leur variance :

4.5.1. Effectifs globaux du trimestre :

Le nombre d'individus débarqués pour une espèce au cours d'un trimestre donné sera finalement estimé par :

$$\boxed{\hat{N}_{tot} = \hat{X}_+ = \sum_h^H \frac{N_h}{n_h}\left(\sum_j^{n_h}\frac{M_{hj}}{m_{hj}}\sum_i^{m_{hj}} X_{hji}\right)}$$

La variance de cet estimateur pourra être estimée par :

$$V\hat{a}r(\hat{X}_+) = \sum_h^H \frac{N_h(N_h-n_h)}{n_h(n_h-1)}\sum_j^{n_h}(\hat{X}_{hj}-\hat{X}_h)^2$$
$$+\sum_h^H\left(\frac{N_h}{n_h}\right)\sum_j^{n_h}\frac{M_{hj}(M_{hj}-m_{hj})}{m_{hj}(m_{hj}-1)}\left(\sum_{i=1}^{m_{hj}}(\hat{X}_{hji}-\hat{X}_{hj})^2\right)$$

Cette expression fait apparaître deux composantes résultant, l'une de la variance entre les jours, l'autre de la variance à l'intérieur des jours.

4.5.2. Effectifs aux longueurs du trimestre :

Pour l'espèce et le trimestre considérés, nous estimerons les effectifs de longueur l débarqués par :

$$\hat{N}_l = \hat{X}_{l+} = \sum_h^H \frac{N_h}{n_h}\left(\sum_{j=1}^{n_h} \frac{M_{hj}}{m_{hj}} \sum_i^{m_{hj}} \hat{X}_{lhji}\right)$$

Et la variance de cet estimateur par :

$$V\hat{a}r(\hat{N}_l) = \sum_{h=1}^{H} \{\frac{(N_h - n_h)N_h}{n_h(n_h-1)} \sum_j^{n_h} (\hat{X}_{lhj} - \hat{X}_{lh})^2\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{N_h}{n_h}\right) \sum_j^{n_h} \frac{M_{hj}(M_{hj} - m_{hj})}{(m_{hj}-1)m_{hj}} \sum_{i=1}^{m_{hj}} (\hat{X}_{lhji} - \hat{X}_{lhj})^2\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{1-f_h}{f_h}\right) \sum_j^{n_h} \left(\frac{M_{hj}}{m_{hj}}\right) \sum_{i=1}^{m_{hj}} \frac{\hat{X}_{lhji}(\hat{X}_{hji} - \hat{X}_{lhji})}{\hat{X}_{hji}}\}$$

Cette fois, trois sources de variance interviennent, d'abord la composante "interjours" dûe au premier niveau d'échantillonnage, ensuite la composante "intercaisses" dûe au deuxième niveau et enfin la composante "intracaisses" résultant de l'EAS dans chaque caisse.

Par ailleurs, la covariance entre deux estimateurs pour des longueurs différentes sera évaluée par l'expression :

$$C\hat{o}v(\hat{N}_{l1}, \hat{N}_{l2}) = \sum_{h=1}^{H} \{\frac{(N_h - n_h)N_h}{n_h(n_h-1)} \sum_j^{n_h} (\hat{X}_{l1hj} - \hat{X}_{l1h})(\hat{X}_{l2hj} - \hat{X}_{l2h})\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{N_h}{n_h}\right) \sum_j^{n_h} \frac{M_{hj}(M_{hj} - m_{hj})}{(m_{hj}-1)m_{hj}} \sum_{i=1}^{m_{hj}} (\hat{X}_{l1hji} - \hat{X}_{l1hj})(\hat{X}_{l2hji} - \hat{X}_{l2hj})\}$$
$$- \sum_{h=1}^{H} \{\left(\frac{1-f_h}{f_h}\right) \sum_j^{n_h} \left(\frac{M_{hj}}{m_{hj}}\right) \sum_{i=1}^{m_{hj}} \frac{\hat{X}_{l1hji}\hat{X}_{l2hji}}{\hat{X}_{hji}}\}$$

4.5.3. Effectifs aux âges du trimestre :

Les effectifs aux âges du trimestre seront estimés par :

$$\hat{N}_a = \hat{X}_{a+} = \sum_h^H \frac{N_h}{n_h}\left(\sum_{j=1}^{n_h} \frac{M_{hj}}{m_{hj}} \sum_i^{m_{hj}} \hat{X}_{ahji}\right)$$

Et la variance :

$$V\hat{a}r(\hat{N}_a) = \sum_{h=1}^{H} \{\frac{(N_h - n_h)N_h}{n_h(n_h - 1)} \sum_j^{n_h} (\hat{X}_{ahj} - \hat{X}_{ah})^2\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{N_h}{n_h}\right) \sum_j^{n_h} \frac{M_{hj}(M_{hj} - m_{hj})}{(m_{hj} - 1)m_{hj}} \sum_{i=1}^{m_{hj}} (\hat{X}_{ahji} - \hat{X}_{ahj})^2\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{1 - f_h f'_h}{f_h f'_h}\right) \sum_j^{n_h} \left(\frac{M_{hj}}{m_{hj}}\right) \sum_{i=1}^{m_{hj}} \frac{\hat{X}_{ahji}(\hat{X}_{hji} - \hat{X}_{ahji})}{\hat{X}_{hji}}\}$$

Enfin, la covariance entre deux estimateurs concernant des âges différents sera évaluée par l'expression :

$$C\hat{o}v(\hat{N}_{a1}, \hat{N}_{a2}) = \sum_{h=1}^{H} \{\frac{(N_h - n_h)N_h}{n_h(n_h - 1)} \sum_j^{n_h} (\hat{X}_{a1hj} - \hat{X}_{a1h})(\hat{X}_{a2hj} - \hat{X}_{a2h})\}$$
$$+ \sum_{h=1}^{H} \{\left(\frac{N_h}{n_h}\right) \sum_j^{n_h} \frac{M_{hj}(M_{hj} - m_{hj})}{(m_{hj} - 1)m_{hj}} \sum_{i=1}^{m_{hj}} (\hat{X}_{a1hji} - \hat{X}_{a1hj})(\hat{X}_{a2hji} - \hat{X}_{a2hj})\}$$
$$- \sum_{h=1}^{H} \{\left(\frac{1 - f_h f'_h}{f_h f'_h}\right) \sum_j^{n_h} \left(\frac{M_{hj}}{m_{hj}}\right) \sum_{i=1}^{m_{hj}} \frac{\hat{X}_{a1hji}\hat{X}_{a2hji}}{\hat{X}_{hji}}\}$$

5. APPLICATION AUX DEBARQUEMENTS DE MORUE A LORIENT :

Nous avons appliqué les calculs à l'exemple des débarquements de morue du deuxième trimestre 1988 sur la criée de Lorient. Les échantillonnages effectués sont très précisément décrits dans des feuilles de relevés dont un exemple est reproduit sur les figures 1 et 2. Toutes les valeurs nécessaires relatives aux **échantillons** sont donc disponibles. Par ailleurs, les débarquements sont comptabilisés au niveau de la criée et stockés dans une base de données. Ainsi, les quantités nécessaires relatives à la **population** sont également connues. Nous pouvons alors, à partir de ces données, obtenir des estimations des moyennes et des variances des effectifs aux âges exprimées au §4.5.3. Les moyennes et les variances assorties des coefficients de variation sont reportées dans le tableau 2.

AGE	MOYENNE	VARIANCE	COEF. DE VAR.
1	37908	14670	38.7%
2	434226	9549.E06	22.5%
3	14898	9615086	20.8%
4	8912	3526022	21.1%
5	2397	1549617	51.9%
6	2603	1911399	53.1%
7	0	0	–
8	24	429.5	86.4%
9+	541	140880	69.4%
7+	565	134899	65.0%

Tableau 2. Estimations de la moyenne et de la variance pour les débarquements du 2ème trimestre 1988 à Lorient.

Au cours du deuxième trimestre 1988, les débarquements sont constitués de poissons âgés de 1 à 9 ans. Les poissons d'âge supérieur à 9 ans sont probablement rangés dans le groupe d'âge 9+. Aucun animal de 7 ans n'a été échantillonné, l'estimation correspondante est donc nulle, mais elle est certainement inférieure à la réalité. Les poissons de plus de 7 ans sont des individus rares dans les débarquements. Une estimation précise de leur nombre nécessiterait donc un effort d'échantillonnage plus

Port : Lorient CABILLAUD Secteur : 7 F-G

Trim : 2e

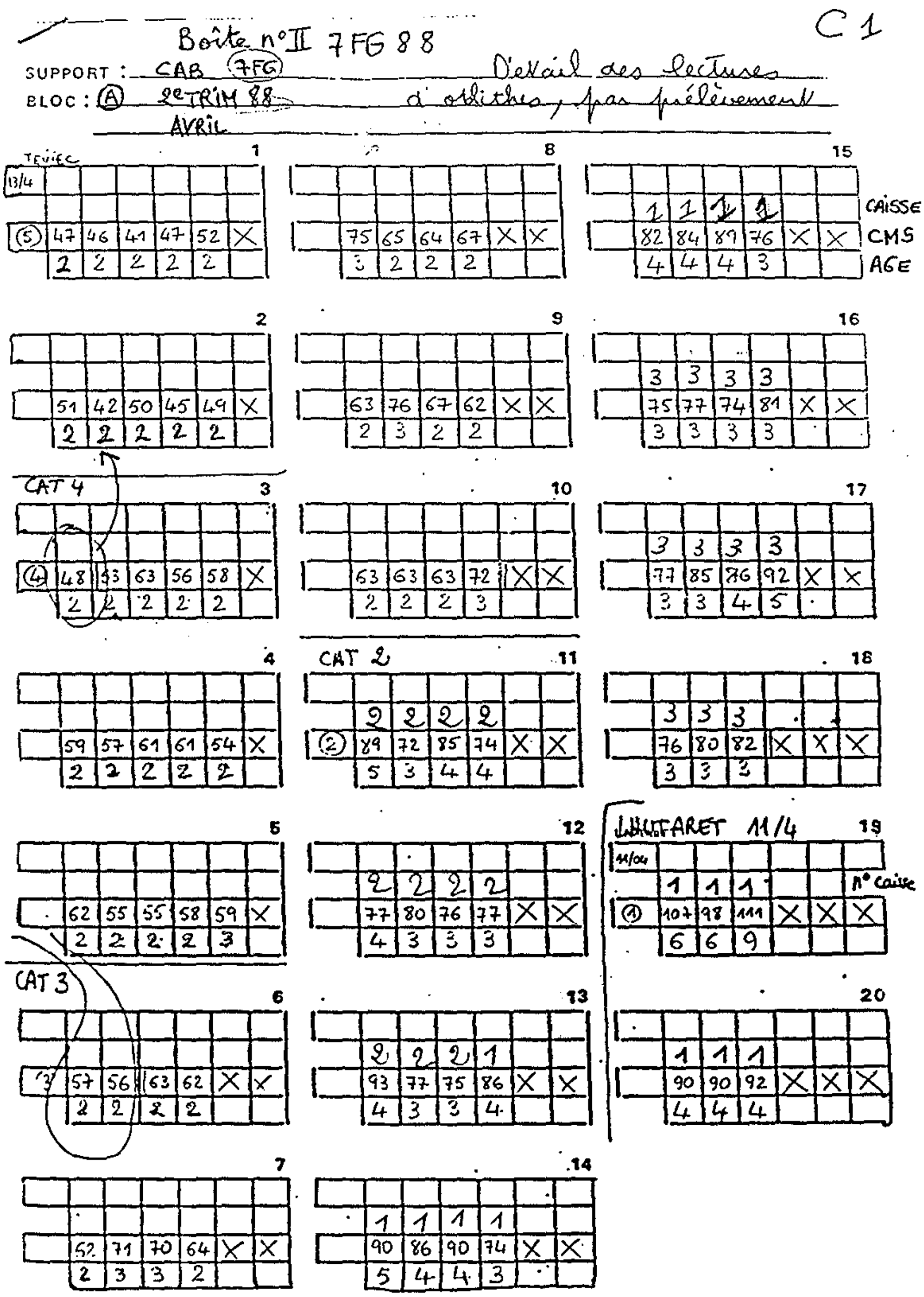
Mois : MAI Navire : D'ARTAGNAN Sous-secteurs :

Date : 09/05/88 Type : Puis. : Ep / sous-sect :

Arm^t : Pot. :

Cat./loc.	5 mini			5			4			3			2				1				
Apports				25 c			30 c			10 c			10 c				2 c				
Ech.	0 c			1/2 c			1 c			1c			3 c				2c				
	20			20			30			50			60				70				
	1			1			1			1			1				1				
	2			2			2			2			2				2				
	3			3			3			3			3				3				
	4			4			4			4			4				4				
	5			5			5			5			5				5				
	6			6			6			6			6				6				
	7			7			7			7			7				7				
	8			8			8			8			8				8				
	9			9			9			9	3		9				9				
	30			30			40			60	1		70				80				
	1			1			1			1	4		1		1	1	1				
	2			2			2			2	4		2	1		1	2				
	3			3			3			3	2		3				3				
	4			4			4			4	1		4				4				
	5			5			5			5	2		5				5				
	6			6			6			6	2		6	4	2	1	6				
	7			7			7			7			7	2	2		7				
	8			8			8	2		8			8		2		8				
	9			9			9	3		9			9		1	1	9	1			
	40			40			50	2		70			80		1	1	90		1		
	1			1			1	2		1	1		1			1	1	1			
	2			2			2	7		2	1		2	2			2	2			
	3			3	1		3	4		3			3			1	3	1			
	4			4			4	4		4	1		4	1	2	3	4	1			
	5			5	4		5	5		5			5	1		1	5		1		
	6			6	3		6	1		6			6				6				
	7			7	7		7	2		7			7	-		1	7		1		
	8			8	6		8	1		8			8	1	1		8		2		
	9			9	6		9			9			9				9				
	50			50			60			80			90				100				
	1			1			1			1			1				1	1	1		
	2			2			2			2			2				2				
	3			3			3			3			3				3				
	4			4			4			4			4				4				
	5			5			5			5			5				5				
	6			6			6			6			6				6				
	7			7			7			7			7				7		1		
	8			8			8			8			8				8				
	9			9			9			9			9				9				
	60			60			70			90			100				110				
	1			1			1			1			1				1				
	2			2			2			2			2				2	1			
	3			3			3			3			3				3				
	4			4			4			4			4				4				
	5			5			5			5			5				5				
	6			6			6			6			6				6				
	7			7			7			7			7				7				
	8			8			8			8			8				8				
	9			9			9			9			9				9				
	70			70			80			100			110				120				

Figure 1. Exemple de fiche de relevé des mensurations.


Boîte n°II 7FG 88

C 1

SUPPORT : CAB 7FG — Détail des lectures d'otolithes, par prélèvement

BLOC : A 2eTRIM 88 AVRIL

TEVIEC 13/4

		CAISSE	CMS	AGE
1	5		47 46 41 47 52 X	2 2 2 2 2
2			51 42 50 45 49 X	2 2 2 2 2
3 (CAT 4)	4		48 53 63 56 58 X	2 2 2 2 2
4			59 57 61 61 54 X	2 2 2 2 2
5			62 55 55 58 59 X	2 2 2 2 3
6 (CAT 3)	3		57 56 63 62 X X	2 2 2 2
7			62 71 70 64 X X	2 3 3 2
8			75 65 64 67 X X	3 2 2 2
9			63 76 67 62 X X	2 3 2 2
10			63 63 63 72 X X	2 2 2 3
11 (CAT 2)	2	2 2 2 2	89 72 85 74 X X	5 3 4 4
12		2 2 2 2	77 80 76 77 X X	4 3 3 3
13		2 2 2 1	93 77 75 86 X X	4 3 3 4
14		1 1 1 1	90 86 90 74 X X	5 4 4 3
15		1 1 1 1	82 84 89 76 X X	4 4 4 3
16		3 3 3 3	75 77 74 81 X X	3 3 3 3
17		3 3 3 3	77 85 76 92 X X	3 3 4 5
18		3 3 3	76 80 82 X X X	3 3 3
19 (LAUTARET 11/4, 11/04, N° Caisse)	1	1 1 1	107 98 111 X X X	6 6 9
20		1 1 1	90 90 92 X X X	4 4 4

Figure 2. Exemple de fiche de lecture d'âge.

important ou une autre procédure. Par la suite, les classes d'âge supérieures ou égales à 7 ans seront regroupées en une seule, comme c'est le cas pour la plupart des analyses de cohorte (existence d'un groupe plus, ici 7+). Du fait de la corrélation entre les effectifs estimés pour les vieux poissons, l'estimation de l'effectif du groupe 7+ possède une variance inférieure. Si l'on excepte les animaux de 5 ans et plus, les variances ne sont pas très élevées dans l'ensemble. Ces poissons âgés sont de toutes façons assez rares et participent pour une faible part seulement au bilan pondéral des captures et du stock. La sensibilité des analyses de cohortes aux captures d'animaux âgés est en effet relativement faible [6].
Notre démarche permet également d'estimer les covariances entre estimateurs. Ces valeurs sont regroupées sous forme de coefficients de variation et de covariation dans le tableau 3.

AGE	1	2	3	4	5	6	7+
1	38.7	0	0	0	0	0	0
2		22.5	3.27	1.22	0	0	0
3			20.8	13.24	0	0	0
4				21.1	10.47	20.33	14.23
5					51.9	19.84	-11.21
6						53.1	25.19
7+							65.

Tableau 3. Coefficients de variation et covariation en pourcentages pour les estimations précédentes dans le cas d'un groupe plus.

On remarque que les covariances sont généralement faibles, comparées aux variances. De plus, elles sont essentiellement positives. Pour les âges qui ne sont jamais échantillonnés le même jour, cette covariance est nulle et les estimateurs sont mêmes indépendants. En ce qui concerne les âges pour lesquels on a trouvé au moins une fois plus d'un individu dans une même caisse, on s'attendrait à des covariances négatives, comme pour toute loi multinomiale. Or, ce n'est pas souvent le cas du fait de l'existence de plusieurs composantes pour la variance. La répartition de la variance selon ses composantes est reportée dans le tableau 4.

AGE	INTERJOURS	INTERCAISSES	INTRACAISSES
1	100	0	0
2	99.99	1E-04	0
3	87.4	11.8	8E-03
4	70.6	26.9	2.5
5	74.9	21.7	3.4
6	93.9	3.5	2.5
7	-	-	-
8	100	0	0
9	8	91.9	0.1

Tableau 4. Composantes de la variance en pourcentages.

La composante intracaisses apparaît souvent négligeable voire nulle. Ceci peut résulter de deux situations différentes : soit la fréquence d'échantillonnage est égale à 1, soit tous les animaux de chaque caisse étaient du même âge.
Dans le premier cas, l'estimation dans chaque caisse correspond en fait à la valeur réelle et sa variance est donc nulle. Ceci se produit pour les poissons de plus de 5 ans qui ne se trouvent que dans la catégorie commerciale 1 pour laquelle tous les poissons de chaque caisse sont mesurés et leur âge déterminé. Néanmoins, les valeurs du tableau ne sont pas nulles car nous avons tenu compte des otolithes illisibles ou ratés qui diminuent le taux réel d'échantillonnage.
Dans le deuxième cas, la distribution multinomiale n'existe pas, chaque tirage donnant un individu du même âge. C'est le cas des individus d'âge 1 qui ne se trouvent que dans la catégorie commerciale 5 mini. Les poissons de cette catégorie sont en fait uniquement mesurés. Cependant, les histogrammes en longueurs des catégories 5 et 5 mini apparaissent disjoints dans les données disponibles. Nous avons par conséquent supposé que tous les animaux classés en 5 mini avaient un an. Une caisse de cette catégorie ne contient donc que des poissons d'un an.
Dans les autres cas, on remarque que la composante intracaisses est négligeable soit parce que le taux d'échantillonnage est élevé, soit parce qu'une des classes d'âge est toujours prépondérante par rapport aux autres (voir la formule du §4.5.3).
En ce qui concerne la composante intercaisses, elle semble augmenter avec l'âge du poisson. Les estimations pour les âges 7 et 8 doivent en effet être considérées à part puisque ce trimestre-là, ne furent trouvés qu'un seul individu de 8 ans et aucun de 7 ans. Cette composante est probablement sous-estimée car dans certaines catégories, une seule caisse est prélevée par jour d'échantillonnage. Dans ces conditions, la variabilité intercaisses du jour ne peut être estimée et une valeur nulle lui est affectée. Dans la mesure où plusieurs caisses sont examinées pour la catégorie 1 pour chaque jour où elle est échantillonnée, cette composante est donc non nulle et même élevée pour l'âge 9, ceci est dû à la rareté de ces individus. Pour les poissons d'âge 3 à 5, le

taux journalier d'échantillonnage des caisses diminue et par suite augmente la composante intercaisses.
Quant à la composante interjours, elle est prépondérante dans la plupart des cas et reflète certainement la diversité des débarquements selon les jours. Effectivement, les données montrent que la composition des captures est très variable en quantité et en qualité selon les jours. Néanmoins, le taux d'échantillonnage des jours est toujours assez faible si l'on considère chaque catégorie. Certains jours ont donné lieu à l'échantillonnage d'une seule catégorie, par exemple 1 ou 5 mini.

Conclusion :

Après examen de ces résultats bruts et suite à l'étude de la propagation de ces résultats dans les modèles [6], il ressort que l'on peut diminuer l'effort d'échantillonnage pour les âges supérieurs à 5 ans, voire à 4, c'est à dire pour la catégorie 1 de cette espèce. A l'opposé, pour certains âges, la composante intercaisse de la variance ne peut être estimée. Il serait préférable de diminuer le taux d'échantillonnage à l'intérieur d'une caisse et en contrepartie, d'échantillonner plus d'une caisse le même jour. Enfin, pour diminuer la composante interjours, il serait souhaitable d'augmenter le nombre de jours à échantillonner. A coût d'échantillonnage égal, il semble que l'on pourrait reporter de l'effort de la catégorie 1 vers les catégories importantes pour le bilan du stock, c'est à dire les catégories commerciales 3,4 et 5. La catégorie 5 mini ne pose pas de problème dans la mesure où elle se recouvre très peu avec la catégorie 5. Cet effort tendrait plutôt à augmenter le nombre de jours qu'à approfondir les prélèvements journaliers. Ces résultats dépendent bien entendu de la méthode et de l'échantillon utilisés. Cependant, les calculs n'ont pas nécessité beaucoup d'hypothèses simplificatrices et leur caractère analytique écarte tout artefact numérique.
Du point de vue pratique, l'obtention d'estimateurs et de variances sur ces estimateurs est assez simple. Les calculs n'ont pas été effectués pour les longueurs du fait du grand nombre d'estimateurs à produire. Pour toute utilisation ultérieure, l'application numérique serait facilitée par l'existence d'un programme de saisie et de traitement des données nécessaires.

Remerciements :

Ce travail n'aurait pas été possible sans la disponibilité et la grande patience de Mrs A. Charuau et J.L. Avrilla du laboratoire IFREMER de Lorient. Je tiens de plus à remercier Mrs R. Chevalier,J. Gueguen et A. Maucorps pour leur relecture attentive du manuscrit.

BIBLIOGRAPHIE :

[1] Chevalier R., 1985, Echantillonnage des débarquements en criée pour l'étude des structures démographiques, *Rev. Trav. Inst. Pêches Marit.*, **49** (3 et 4): 95-110.

[2] Chevalier R., B. Mesnil, 1986, Echantillonnage en criée pour l'estimation des structures démographiques, DRV-86.006-RH/NANTES.

[3] Chevalier R., E. Gondeaux, C. Koutsikopoulos, 1988, Séminaire d'initiation à la théorie et à la pratique des sondages, Rapport DRV-88.027-RH/NANTES.

[4] Cochran W.G., 1977, Sampling techniques, 3ème édition, Wiley.

[5] Etude d'une gestion optimale des pêcheries de langoustine et de poissons démersaux en mer celtique, 1989, Rapports internes de la Direction des Ressources Vivantes de l'IFREMER, Tomes 1,2 et 3, DRV-89.009-RH/LORIENT.

[6] Pelletier D., 1990, Sensitivity and variance estimators for virtual population analysis and the equilibrium yield per recruit model, *Aquat. Liv. Res.,* **3** (1) : 1-12.

[7] Tassi P., 1985, Méthodes statistiques, Economica.

ANNEXE :

Il est nécessaire de rappeler certains résultats classiques en échantillonnage auxquels il est fait appel plusieurs fois dans l'exposé. Nous nous contenterons des formules finales. Tous les détails du raisonnement et des calculs se trouvent dans [3] ou [4].

–I–

CAS D'UN ECHANTILLONNAGE ALEATOIRE SIMPLE :

Soit une population d'effectif total M dans laquelle on tire au hasard et sans remise m individus sur lesquels on mesure une variable X. Si X_i est la valeur de X pour l'individu i, on estime la moyenne de X dans la population par :

$$\hat{X} = \frac{1}{n}\sum_{i=1}^{m} X_i$$

Et la variance de cet estimateur est :

$$Var(\hat{X}) = \left(\frac{M-m}{M}\right)\frac{S^2}{m}$$

où :

$$S^2 = \frac{\sum_{i=1}^{M}(X_i - X)^2}{M-1}$$

Que l'on estime grâce aux données de l'échantillon par :

$$s^2 = \frac{\sum_{i=1}^{m}(X_i - X)^2}{m-1}$$

L'expression de l'estimation de la variance est donc finalement :

$$V\hat{a}r(\hat{X}) = \left(\frac{M-m}{M}\right)\frac{s^2}{m}$$

Que l'on exprime souvent, en notant f le taux d'échantillonnage m/M, par :

$$V\hat{a}r(\hat{X}) = (1-f)\frac{s^2}{m}$$

La covariance entre deux variables dont on estime les moyennes X et Y sur le même échantillon s'obtient de manière similaire et on peut l'estimer par :

$$C\hat{o}v(\hat{X},\hat{Y}) = \frac{(1-f)}{m}\frac{\sum_{i=1}^{m}(X_i - X)(Y_i - Y)}{m-1}$$

-II-

ECHANTILLONNAGE STRATIFIE :

Si la population à laquelle nous nous référons est partitionnée en classes dans *chacune* desquelles on effectue des tirages aléatoires simples *indépendants*, on dit que l'échantillonnage est aléatoire simple stratifié (EAS stratifié).

Soient H strates de tailles (i.e. d'effectifs) respectives M_h. On appelle poids de la strate la quantité suivante :

$$w_h = \frac{M_h}{M}$$

où M est la taille de la population soit la somme des tailles des strates.

On tire m_h poissons dans chaque strate. Supposons qu'on mesure une variable X sur chacun des individus tirés. Si X_{hi} est la valeur de X pour l'individu i de la strate h, on estimera la moyenne de X sur la population par :

$$\hat{X} = \sum_{h=1}^{H} w_h \overline{X}_h$$

où :

$$\overline{X}_h = \frac{1}{m_h} \sum_{i=1}^{m_h} X_i$$

Quant à la variance de cet estimateur, elle s'exprime simplement du fait de l'indépendance des strates par :

$$Var(\hat{X}) = \sum_{h=1}^{H} w_h^2 Var(\overline{X}_h)$$

Or, la variance à l'intérieur d'une strate est celle d'un EAS de taille m_h. On obtient donc la formule :

$$Var(\hat{X}) = \sum_{h=1}^{H} w_h^2 (1 - f_h) \frac{s_h^2}{m_h}$$

En particulier, si le taux d'échantillonnage est identique dans chaque strate, l'expression se simplifie en :

$$Var(\hat{X}) = (1 - f) \sum_{h=1}^{H} w_h^2 \frac{s_h^2}{m_h}$$

-III-

ECHANTILLONNAGE A DEUX NIVEAUX :

La population est cette fois divisée en N unités qu'on appelle primaires, chaque unité i se compose elle-même de M_i unités dites secondaires. L'échantillonnage consiste à prélever n unités primaires puis à pratiquer un deuxième tirage, par exemple un EAS ou un EAS stratifié sur chacune des unités primaires sélectionnées. Si l'on mesure une variable X sur chacun des individus tirés, dans le cas d'un EAS de taille m_i sur chaque unité primaire i, on estimera la moyenne de X dans la population par :

$$\hat{X} = \frac{1}{n}\sum_{i=1}^{n}\overline{X}_i$$

où :

$$\overline{X}_i = \frac{1}{m_i}\sum_{j=1}^{m_h} X_x$$

La variance de cet estimateur est la somme de deux termes :

$$Var(\hat{X}) = Var_1 E_2(\hat{X}) + E_1 Var_2(\overline{X}_i)$$

En effet, le premier terme correspond à la variance inter-unités primaires tandis que le second décrit la variance intra-unités primaires.

Supposons qu'on estime également la moyenne d'une autre variable Y à partir de l'échantillon. La covariance entre les deux estimateurs vaudra :

$$Cov(\hat{X},\hat{Y}) = Cov_1(E_2(\hat{X}),E_2(\hat{Y})) + E_1(Cov_2(\hat{X},\hat{Y}))$$

On peut encore estimer ces deux expressions par :

$$V\hat{a}r(\hat{X}) = \frac{N-n}{Nn}\frac{\sum_{i}^{n}(\hat{X}_i-\hat{X})^2}{n-1} + \frac{\sum_{i}^{n}Var_2(\overline{X}_i)}{nN}$$

$$C\hat{o}v(\hat{X},\hat{Y}) = \frac{N-n}{Nn}\frac{\sum_{i}^{n}(\hat{X}_i-\hat{X})(\hat{Y}_i-\hat{Y})}{n-1} + \frac{\sum_{i}^{n}Cov_2(\overline{X}_i,\overline{Y}_i)}{nN}$$

Selon la nature du deuxième tirage, les expressions de $Var_2(\overline{X}_i)$ et de $Var_2(\overline{Y}_i)$ seront différentes. Ce seront par exemple celles d'un EAS ou d'un échantillonnage aléatoire stratifié.

-IV-

CARACTERISTIQUES D'UNE LOI MULTINOMIALE :

La loi multinomiale apparaît fréquemment dans les problèmes d'échantillonnage dès lors que l'on s'intéresse à des individus répartis en classes disjointes identifiables après tirage. Elle est en fait une généralisation de la loi binomiale [7].

Considérons une population composée d'éléments de L types, en proportions respectives $p_l, l = 1, L$. On tire dans cette population de façon équiprobable et indépendante, un échantillon de taille n et on écrit le vecteur :

$$(D_1, D_2, \ldots, D_l, \ldots D_L)$$

composé des nombres d'éléments de chaque type qu'on trouvera dans l'échantillon. On dit que ce vecteur suit une loi multinomiale de paramètres :

$$(n, p_1, p_2, \ldots, p_l, \ldots p_L)$$

Cette loi possède les propriétés suivantes :

$$E(D_l) = n p_l$$

$$Var(D_l) = n p_l (1 - p_l)$$

et pour $l_1 \neq l_2$:

$$Cov(D_{l1}, D_{l2}) = -n p_{l1} p_{l2}$$

Supposons maintenant qu'on a effectué un EAS de taux f dans la population et qu'on veuille estimer la moyenne dans la population du nombre d'individus de type l. On estimera cette quantité par :

$$\hat{X}_l = D_l / f$$

où D_l prendra la valeur trouvée dans l'échantillon.

Notons D la taille de l'échantillon. On a :

$$D = fX = f \sum_{l=1}^{L} X_l / f$$

On en déduit alors la variance de cet estimateur :

$$Var(\hat{X}_l) = \frac{(1-f)}{D} Var(D_l)$$

Soit, après substitution :

$$Var(\hat{X}_l) = \frac{(1-f)}{f} p_l (1-p_l) X$$

De la même manière, on obtient le terme de covariance entre les deux estimateurs $\hat{X}_{l1}$ et $\hat{X}_{l2}$:

$$Cov(\hat{X}_{l1}, \hat{X}_{l2}) = -\frac{(1-f)}{f} p_{l1} p_{l2} X$$