



Pêches et Océans
Canada

Fisheries and Oceans
Canada

Sciences des écosystèmes
et des océans

Ecosystems and
Oceans Science

Secrétariat canadien de consultation scientifique (SCCS)

Document de recherche 2019/013

Région de la capitale nationale

Cadre unifié pour l'évaluation statistique des programmes de surveillance des pêches

Jacques Allard¹ et Hugues P. Benoît²

¹ Atlantic Statistical Analysis Inc.
Dieppe (N.-B.) E1A 4Z7

² Pêches et Océans Canada
Institut Maurice-Lamontagne
Mont-Joli (Québec) G5H 3Z4

14. ANNEXE II - GLOSSAIRE

Attendu (voir prévu): Utilisé au sens statistique de « valeur attendue »

Biais (opérationnel): Différence systématique entre la valeur attendue d'un estimateur et la valeur réelle du paramètre estimé en raison des caractéristiques opérationnelles du protocole. Applicable aux échantillons aléatoires et aux recensements.

Biais (statistique): La différence entre la valeur attendue d'un estimateur et la valeur réelle du paramètre estimé en raison des propriétés mathématiques de l'estimateur. Le biais est une valeur signée: un biais négatif correspond à une sous-estimation, un biais positif à une surestimation. Applicable uniquement aux échantillons aléatoires.

Bootstrap: Lorsqu'une estimation est obtenue à partir d'un échantillon aléatoire, souvent on ne peut obtenir la distribution de probabilité de l'estimation à partir de théorèmes mathématiques. Dans la plupart des cas, la distribution de probabilité de l'estimation peut être obtenue, au moins approximativement, par une méthode informatique intensive appelée « bootstrap », dans laquelle les données sont rééchantillonnées comme si leur ensemble constituait la population. Une fois la distribution de probabilité d'une estimation obtenue, on peut obtenir son erreur-type et son biais.

Couverture: Dans l'échantillonnage, le sous-ensemble de la population qui est inclus dans le processus de sélection de l'échantillon. Dans un recensement, le sous-ensemble de la population pour lequel une observation est faite ou une mesure prise. *EXEMPLE* de gestion des pêches: Les échantillons aléatoires ou les recensements de pêcheurs sportifs ne peuvent inclure que les pêcheurs titulaires d'un permis, à l'exclusion des pêcheurs non titulaires d'un permis.

Données d'observateurs indépendants: Mesures recueillies par une personne ou une technologie spécifiquement chargée d'observer les activités de pêche et d'en faire rapport sans lien de dépendance avec l'industrie ou la collectivité de la pêche, comme les observateurs en mer et à quai, les caméras embarquées et les systèmes de surveillance des navires.

Données provenant des utilisateurs de la ressource: Mesures prises ou déclarées par l'industrie ou la collectivité de la pêche, y compris les pêcheurs, le personnel de l'usine et les acheteurs. Les registres du journal de bord, les bordereaux d'achat et les réponses aux sondages auprès des pêcheurs sportifs sont des exemples de données provenant des utilisateurs de la ressource.

Erreur (opérationnelle): Différence prévue entre une estimation et la valeur estimée dans un sens positif ou négatif. *EXEMPLE*: Erreur due à une erreur d'étiquetage d'une prise.

Erreur (statistique): Différence prévue entre une estimation et la valeur estimée dans un sens positif ou négatif en raison du caractère aléatoire d'un protocole d'échantillonnage, le plus souvent décrit par l'erreur-type. Ne s'applique qu'à l'échantillonnage aléatoire.

Erreur-type (d'un estimateur): La description la plus populaire de la variation d'un estimateur causée par le caractère aléatoire d'un protocole d'échantillonnage. Notation typique: $ET, \sigma_{\hat{\theta}}$.

Erreur-type (ET): La norme d'un estimateur est une mesure de l'erreur due uniquement au caractère aléatoire du protocole d'échantillonnage.

Estimateur: Méthode permettant d'obtenir une estimation à partir d'un échantillon à partir de données provenant d'un protocole d'échantillonnage ou d'un recensement. *EXEMPLES*: La moyenne d'échantillon d'un échantillon obtenu par échantillonnage aléatoire simple est un estimateur de la moyenne de la population. Une moyenne tronquée à 5 % (les 5 % les plus

petits et les 5 % les plus grands des observations étant supprimés avant le calcul de la moyenne) est également un estimateur de la moyenne de la population.

Estimation : Quantité dérivée d'un échantillon ou d'un recensement utilisée comme approximation du paramètre. **EXEMPLE** de gestion des pêches : Utilisation du produit [prise par unité d'effort PUE (à partir de l'échantillonnage)] fois [effort (à partir des journaux de bord)] pour estimer la prise totale. Notation typique : $\hat{\theta}$ pour le paramètre θ . Le terme estimation peut aussi désigné le processus d'obtention d'une estimation.

L'**exactitude** fait référence à l'absence d'erreur systématique, c'est-à-dire une erreur qui a une valeur attendue positive ou négative. L'exactitude est l'inverse du *biais*.

Fiabilité : Terme utilisé ici pour exprimer la qualité d'un processus d'estimation en tant qu'outil permettant d'atteindre les objectifs de la gestion des pêches. La qualité des estimateurs est un déterminant important de la fiabilité.

Imputation : Processus par lequel une valeur manquante est remplacée par une valeur supposée proche de la valeur manquante. Il existe plusieurs méthodes statistiques pour effectuer l'imputation. Pêches. **EXEMPLE** : Utilisation de la PUE des rivières voisines pour estimer les captures de l'effort de pêche d'une rivière pour laquelle les valeurs de la PUE ne sont pas disponibles.

Paramètre : La quantité à estimer. Exemples de gestion des pêches : prise totale d'une espèce cible, proportion de crabe blanc dans la prise, prise de morue-lingue par unité d'effort. Notation typique : θ .

Population cible : Ensemble de toutes les unités individuelles qui doivent être observées (recensement) ou échantillonnées (enquête par sondage). Idéalement, le *cadre* et la population cible seraient identiques. S'ils ne sont pas identiques, il y a sous-dénombrement (le cadre est plus petit que la population cible) ou surdénombrement (le cadre est plus grand que la population cible). Exemples : Toutes les sorties de pêche d'une pêche donnée, tous les pêcheurs récréatifs de saumon atlantique.

Post-traitement : Rajustements des données pour compenser les variations dans la collecte. Exemples : poids des prises éviscérées et non éviscérées; conversion de livres en kilogrammes.

Précision fait référence à la répétabilité/reproductibilité d'une mesure, c'est-à-dire une erreur dont la valeur attendue est égale à zéro. La précision est l'inverse de la *variabilité* ou, en statistique, de l'*erreur-type* et de la *variance*.

Prévu (voir attendu) : Nous réservons le mot « attendu » aux applications statistiques, comme dans « valeur attendue ». Nous utilisons le terme « prévu » pour désigner l'acte de prévoir ou de prédire une valeur ou un événement. **Processus d'estimation des paramètres** ou **processus d'estimation** : Processus complet commençant par le choix des unités à observer, les observations, la saisie des données, les divers calculs, etc. et se terminant par le calcul de l'estimation.

Protocole d'échantillonnage : Le processus utilisé pour obtenir un échantillon. Exemple de gestion des pêches : La procédure utilisée pour décider quand un observateur en mer doit être déployé et quelles sont les caractéristiques de la sortie de pêche qu'il est censé surveiller et enregistrer.

Qualité : La mesure de la proximité d'une estimation par rapport à la valeur réelle : elle résume l'*exactitude* et la *précision*.

Recensement : Une enquête où l'ensemble de la population est ciblée à des fins d'observation.

Théorème de la limite centrale : Lorsqu'une estimation est obtenue à partir d'un échantillon aléatoire, la distribution de probabilité d'un estimateur peut parfois être obtenue à partir de théorèmes mathématiques. Dans le cas le plus simple, la moyenne d'un échantillon aléatoire simple est utilisée pour estimer la moyenne de la population : dans ce cas, le théorème de la limite centrale indique que la distribution de probabilité de la moyenne de l'échantillon est une distribution normale (aussi appelée gaussienne), du moins si l'échantillon est grand. Il a été prouvé mathématiquement que la distribution de probabilité de nombreuses autres estimations est une distribution normale. Une fois la distribution de probabilité d'un estimateur obtenue, on peut calculer son erreur-type et son biais.

Variance (d'un estimateur) : Le carré de l'erreur-type. Les statisticiens préfèrent calculer avec la variance parce que certaines formules deviennent plus simples et parce que l'« analyse de variance » permet un partitionnement de la variance par source (mais pas de l'erreur-type).

Variation d'échantillonnage (d'un estimateur) : La variation d'un estimateur causée par le caractère aléatoire d'un protocole d'échantillonnage.