



L'épidémiologie pour tous

Exactitude et précision d'une estimation

Rédigée par B. Toma, B. Dufour, J. Rivière et J.J. Bénet ; validée par le bureau de l'AEEMA

Contexte

Dans la vie quotidienne, chacun peut être amené à **estimer** des caractéristiques d'un objet ou d'un sujet, par exemple le prix d'un objet (sans étiquette), la vitesse d'une voiture, la hauteur d'un arbre, le poids et/ou âge d'une personne, etc.

Pour une estimation effectuée de manière quantitative, le résultat peut être formulé :

- de manière unique (par exemple, respectivement : 10 euros, 50 km/h, 8 mètres, 60 kg et 30 ans, etc.) ;
- ou sous forme d'une fourchette (7-15 euros, etc.).

Chaque résultat d'une estimation peut être plus ou moins éloigné de la réalité (donc, plus ou moins **exact**), et plus ou moins **précis** (dans le cadre d'une estimation sous forme d'une plage de valeurs).

Exemple

Ainsi, pour le poids réel d'un sujet de **55 kg**, des estimations sont présentées dans le tableau 1.

Tableau 1. Estimations proposées pour l'estimation du poids d'un sujet

Estimation	Précision	Exactitude
70 kg	Très précise	Inexacte
45-65 kg	Peu précise	Exacte
45-50 kg	Précise	Inexacte
50-60 kg	Précise	Exacte
60-80 kg	Peu précise	Inexacte

On constate donc qu'**exactitude** et **précision** d'une estimation sont deux qualités indépendantes.



Exactitude d'une estimation :
Qualité de l'accord entre l'estimation de la valeur mesurée et la valeur vraie.

Précision d'une estimation :
Dispersion des valeurs d'une estimation, autour de sa valeur moyenne.

Application à l'épidémiologie descriptive à visée quantitative

Facteurs conditionnant la précision d'une estimation

En **épidémiologie descriptive** à visée quantitative, l'estimation de la **prévalence** d'une maladie dans une **population** s'effectue, le plus souvent, sur un **échantillon**, et à l'aide d'un test de **dépistage**.

A partir des résultats obtenus sur l'échantillon, on estime alors la valeur de la prévalence dans la population.

Comme dans la vie quotidienne, le résultat de l'estimation peut être plus ou moins exact et plus ou moins précis, en fonction de caractéristiques du test de dépistage et de l'échantillon utilisé.

- **Test de dépistage**

Le résultat de la prévalence obtenu sur l'échantillon est d'autant plus exact que les qualités intrinsèques du test utilisé sont bonnes (cf. Fiches **Sensibilité d'un**

[test de dépistage](#) et [Spécificité d'un test de dépistage](#)).

- **Echantillon**

➤ La **représentativité** (cf. Fiche [Représentativité d'un échantillon](#)) de l'échantillon est nécessaire pour pouvoir estimer la valeur de la caractéristique étudiée (le plus souvent, la prévalence d'une maladie) au sein de la population, à partir du résultat obtenu sur l'échantillon.



La **représentativité** d'un échantillon **conditionne** l'**exactitude** du résultat.



Un **échantillon non représentatif** de la population de laquelle il est issu **ne permet pas d'estimer correctement**, à partir du résultat qu'il fournit, la valeur probable dans la population.

➤ Pour la **précision du résultat**, obtenu sur un échantillon représentatif et appliqué à la population, la règle est simple :



L'**effectif** (la taille) d'un échantillon **conditionne** la **précision** du résultat.



La **précision relative** du résultat (cf. Fiche [Précision d'une estimation](#)) est **d'autant meilleure** que l'**effectif** de l'échantillon est **important**.

Par exemple, un échantillon de 200 sujets fournit un résultat plus précis qu'un échantillon de 100 sujets et, *a fortiori*, qu'un échantillon de 10 sujets.

Ce lien direct entre l'effectif de l'échantillon et la précision du résultat conduit à déterminer le **nombre de sujets nécessaire** pour constituer un échantillon lors d'une enquête descriptive à visée quantitative, en **fonction** de la **précision relative** désirée et de la **prévalence « attendue »** (ou supposée) de la maladie (cf. Fiche [Nombre de sujets nécessaire pour une enquête descriptive à visée quantitative](#))

Détermination du niveau de précision nécessaire pour une enquête

L'obtention d'un résultat très précis lors d'une enquête descriptive dépend de la taille de l'échantillon utilisé. La contrepartie de cette précision est le coût de l'enquête qui augmente proportionnellement à la taille de l'échantillon. Il est donc important, avant de préparer le protocole qui définit la taille de l'échantillon, de bien réfléchir à la précision qui paraît indispensable : plus ce niveau sera élevé, plus l'enquête sera coûteuse. La question à se poser est de savoir à **quoi servira le résultat de l'enquête** : à quelles(s) décision(s), en fonction de quel(s) seuil(s) de prévalence ?

Détermination du nombre de sujets nécessaire en fonction de la précision souhaitée

Des tableaux permettent de déterminer le nombre de sujets nécessaire en fonction du niveau de précision relative désirée et de la prévalence attendue (tableau 4.2 du livre *Epidémiologie appliquée* pour une population dite « infinie » et tableaux 4.3 et 4.4 lorsque le taux de [sondage](#) est supérieur à 10 %). Sur ces tableaux, on constate que les nombres de sujets sont très élevés pour des prévalences attendues faibles et une précision relative élevée.



Plus la **prévalence attendue** est **faible**, **plus élevé** sera le **nombre** de sujets nécessaire, **pour un même niveau de précision**.

De même, **pour un niveau donné de prévalence attendue**, le **nombre** de sujets nécessaire sera **d'autant plus élevé** que la **précision relative** jugée nécessaire est **importante**.

Un tableur (annexe « [Taille d'échantillon aléatoire simple](#) » du livre *Epidémiologie appliquée*) présent sur le site de l'AEEMA permet également de calculer le nombre de sujets nécessaire.

4. Précision obtenue après l'enquête

Après la réalisation de l'enquête, si le taux de prévalence trouvé est différent de celui attendu, le niveau de précision du résultat n'est pas le même que celui souhaité.

Ainsi, si le pourcentage de **prévalence** trouvé est **plus élevé** que celui qui était attendu, la **précision** obtenue est **meilleure** (et le coût aura été plus élevé que celui correspondant au niveau de précision relative souhaité).

A l'inverse, si le pourcentage de **prévalence** trouvé est **plus faible** que celui attendu, la **précision** du résultat est **plus faible** que celle souhaitée. Dans ce cas, si le

niveau de précision souhaité est indispensable, il convient de procéder à un **complément d'enquête** sur un échantillon représentatif supplémentaire afin d'atteindre le nombre de sujets nécessaire pour un pourcentage de prévalence plus faible que celui qui était attendu.



Prévalence trouvée

Précision

> **attendue**



Meilleure

< **attendue**



Moins bonne

*L'**exactitude** du résultat est **indépendante** de la **précision** du résultat.*

*L'**exactitude** du résultat est conditionnée par la **représentativité** de l'échantillon.*

*La **précision** du résultat est fonction du **nombre de sujets** dans l'échantillon.*

*Pour un **pourcentage trouvé plus élevé** que celui attendu, la **précision** est **meilleure** que souhaitée.*

*Pour un **pourcentage trouvé moins élevé** que celui attendu, la **précision** est **moins bonne** que souhaitée.*



Toma B. et al. : *Epidémiologie appliquée*, 2018, AEEMA éditeur, 614 p.

Annexe : « [Taille d'échantillon aléatoire simple](#) »

Terminologie : des liens hypertextes (en bleu souligné) vous permettent d'accéder aux définitions sur le site de l'AEEMA

Fiches AEEMA : [Sensibilité \(individuelle ou collective\) d'un test de dépistage](#) ;

[Spécificité \(individuelle ou collective\) d'un test de dépistage](#) ;

[Représentativité d'un échantillon](#) ;

[Précision d'une estimation](#) ;

[Nombre de sujets nécessaire pour une enquête à visée quantitative](#)

