



Université Ziane Achour Djelfa
Faculté des Sciences de la Nature et de la Vie
Département de Biologie
Spécialité : biologie et Physiologie Animale



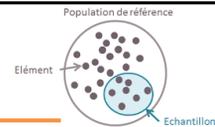
Module: Techniques d'échantillonnage

Calcul de la taille d'un échantillon



Dr. LOUNIS M

Introduction



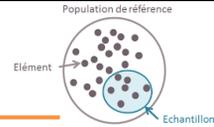
Le calcul de la taille de l'échantillon est une méthode qui va nous permettre d'obtenir le nombre minimum de sujets à inclure dans une étude pour garantir, dans les études descriptives une bonne précision, et dans les études analytiques une bonne puissance.

Étape primordiale dans la réalisation d'un travail de recherche qu'il soit descriptif ou analytique

Une étude épidémiologique réalisée avec un sous-effectif est une perte de temps et de moyens car ne permettra pas d'atteindre l'objectif du travail.

Étape nécessaire est important : seule façon qui permet de garantir la précision ou la puissance adéquate selon l'objectif de l'étude, et d'évaluer la faisabilité en termes de temps et de moyens disponibles.

Introduction

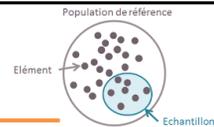


La détermination de la taille de l'échantillon dépend de plusieurs paramètres selon l'objectif de l'étude.

Cette détermination peut se faire grâce à des formules statistiques en se basant sur les résultats escomptés.

Il faut noter que les données de la littérature sont importantes pour orienter le calcul de la taille de l'échantillon.

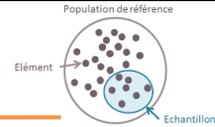
Études descriptives



Afin de calculer la taille de l'échantillon on a besoin de :

- Fixer le risque d'erreur α (souvent α 5 %) ou degré (seuil) de confiance à 95 %
- Déterminer la valeur attendue du paramètre à estimer (généralement proportion ou moyenne) à partir de la littérature ou le résultat d'une étude pilote.
- Fixer le degré de précision désiré qui doit être acceptable.

Méthodes de calcul



Estimation d'une proportion :

$$n = \frac{Z_{\alpha}^2 * p * (1-p)}{e^2}$$

$$n = \frac{Z_{\alpha}^2 * p * q}{e^2}$$

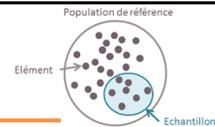
$q=1-p$

n : taille de l'échantillon,

p : proportion attendue,
proportion estimée de la population présentant la caractéristique étudié

e : degré de précision accepté (marge d'erreur que l'on est prêt à accepter en décimales (2% = 0,02).

Z_{α} : valeur critique sur la loi normale centrée réduite pour un risque d'erreur α (pour $\alpha=5\%$, $Z_{\alpha}=1,96$).

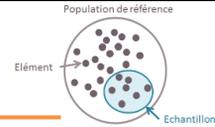


Malheureusement, on ne connaît pas de p .

La seule solution dans ce cas est d'en donner une estimation « à priori »

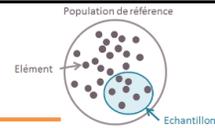
- soit à partir d'une enquête portant sur le même sujet mais réalisée dans le passé,
- soit à partir de conseils d'experts,
- soit réaliser une pré-étude.

Valeur de Z_α



Seuil de confiance souhaité (%)	Valeur de Z_α
80% ($\alpha = 20\%$)	1.28
85% ($\alpha = 15\%$)	1.44
90% ($\alpha = 10\%$)	1.65
95% ($\alpha = 5\%$)	1.96
99% ($\alpha = 1\%$)	2.58
99,99% (0,1%)	3.29

Cas de population finie (échantillon $\geq 1/20$)



Plusieurs formules:

$$n = \frac{Z_\alpha^2 * p * q * N}{(N - 1) * e^2 + Z_\alpha^2 * p * q}$$

$$n = \frac{z^2 * p * (1 - p)}{e^2 + \left(\frac{z^2 * p * (1 - p)}{e^2 N} \right)}$$

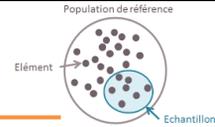
$$n = \frac{p * (1 - p) + \frac{e^2}{Z_\alpha^2}}{\frac{e^2}{Z_\alpha^2} + \frac{p * (1 - p)}{N}}$$

Formules plus simples

$$\hat{n} = \frac{n}{1 + \frac{n-1}{N}}$$

$$n' = \frac{N * n}{N + n}$$

Estimation d'une moyenne



Population infinie:

$$n = \frac{Z_{\alpha}^2 * S^2}{e^2}$$

n: taille de l'échantillon,

S²: variance de des valeurs pour lesquelles on veut calculer la moyenne

e: degré de précision accepté (marge d'erreur que l'on est prêt à accepter)

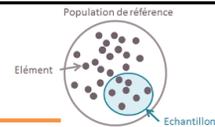
Z_α: valeur critique sur la loi normale centrée réduite pour un risque d'erreur α (pour α=5 %, Z_α =1,96).

Population finie:

$$n = \frac{t_{\alpha}^2 * s^2 * N}{(N - 1) * e^2 + t_{\alpha}^2 * s^2}$$

$$n = N * \frac{1}{1 + (N * e^2 / 4S^2)}$$

Taux de réponse

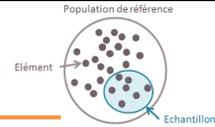


Ajouter le pourcentage de non réponse à la taille de l'échantillon

$$n' = \frac{n}{1 - d}$$

d= proportion de non réponse

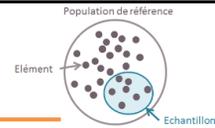
Exercice



On veut estimer la proportion de sujets atteints d'une maladie dans une population. Chez des patients on sait à travers les données de la littérature ou selon l'expérience en tant que spécialiste qu'elle peut être d'environ 13 %.

La proportion attendue (p) est donc de 13 %,
On veut estimer la taille de l'échantillon avec un risque d'erreur alpha de 5 % et une précision de 3 %,

le nombre de sujets nécessaire est:

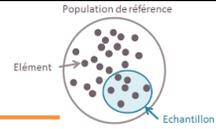


$$n = \frac{1,96^2 \times 0,13 \times (1 - 0,13)}{0,03^2} = 482,8$$

La taille de l'échantillon nécessaire est de 483

Si on suppose que la population est finie et composée de 800 individus

La taille de l'échantillon est?

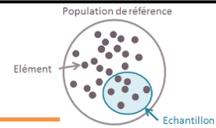


Taux de sondage = 483/800

$$n' = \frac{800 * 483}{800 + 483} = 301,17$$

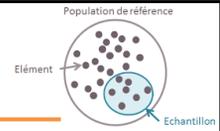
$$n' = \frac{800 *}{1 + \frac{483-1}{800}} = 301,17$$

La taille de l'échantillon est de 302 sujets



On prévoit un taux de non réponse de 20%

Quel serait la taille de l'échantillon:



$$n' = \frac{n}{1 - d}$$

$$n' = \frac{483}{0.8}$$

La taille de l'échantillon est de: 604 sujets