TAILLE D'UN ECHANTILLON ALEATOIRE SIMPLE

I – EN VUE DE L'ESTIMATION D'UNE PROPORTION

La relation entre précision absolue, proportion à estimer et taille de l'échantillon peut s'exprimer par la formule :

$$Pa = 1{,}96 \sqrt{\frac{pq}{n}}$$

avec : Pa = précision absolue (fixée par l'enquêteur)
p = hypothèse sur la proportion à estimer (avec q = 1 – p)
n = taille de l'échantillon (à déterminer)

En élevant au carré les deux membres de l'égalité, on obtient :

$$Pa^2 = 3{,}84 \; \frac{pq}{n}$$

$$\text{D'où :} \quad n = \frac{3{,}84 \; pq}{Pa^2}$$

Ou, en utilisant la précision relative $Pr = \frac{Pa}{p}$ (soit : Pa = Pr x p)

$$\boxed{n = \frac{3{,}84 \; q}{p \; Pr^2}} \quad (1)$$

Cette formule est applicable lorsque $\frac{n}{N}$ < 10 p. cent.

Dans le cas où $\frac{n}{N}$ > 10 p. cent, la taille de l'échantillon n' doit tenir compte de la taille de la population :

$$\boxed{n' = \frac{1}{\frac{1}{n} + \frac{1}{N}}} \quad (2)$$

avec : n' = taille de l'échantillon (à déterminer)
n = taille de l'échantillon calculée avec la formule (1)
N = taille de la population

II – EN VUE DE LA DETECTION D'UN PHENOMENE

1. POPULATION INFINIE (n/N <10 p. cent)

Dans une population atteinte par une maladie de prévalence p, la probabilité de n'avoir aucun animal malade parmi les n animaux tirés au sort est :

$$\alpha = (1 - p)^n \qquad (1)$$

La taille de l'échantillon pour un risque donné α et une prévalence donnée p est déduite de la formule (1) :

$$\text{Log}(\alpha) = \log[(1-p)^n]$$

$$\text{Log}(\alpha) = n\log(1-p)$$

D'où

$$\boxed{n = \frac{\log(\alpha)}{\log(1-p)}} \quad (2)$$

Exemple : α *= 5 p. cent*

P = 1 p. cent

$$n = \frac{log(0{,}05)}{log(1-0{,}01)} = \frac{-1{,}30103}{-0{,}004365} = \mathbf{298}$$

2. POPULATION FINIE (n/N >10 p. cent)

La probabilité de n'avoir aucun malade parmi les n animaux tirés au sort dans une population N est :

$$\alpha = \frac{C_{N-M}^{n}}{C_{N}^{n}} \qquad (3)$$

avec :

- n : taille de l'échantillon
- N : taille de la population
- M : nombre d'unités malades dans la population
- C_{N-M}^{n} : nombre de combinaisons de n éléments parmi les (N – M)

La formule (3) est plus complexe que la formule (2) et sa résolution analytique pour le calcul de n est impossible.

Il existe cependant des solutions approchées pour le calcul de n :

$$\boxed{n = \left[1-(\alpha)^{1/M}\right] \times \left[N - \frac{M}{2}\right] + 1}$$

Exemple :

Si α est fixé à 5 p. cent, si la taille de la population est de 500 et si on suppose que la prévalence est au moins égale à 10 p. cent alors :

N = 500
M = 50 (10 p. cent de 500)
α = 5 p. cent

$$n = \left[1-0{,}05^{1/50}\right] \times \left[500 - \frac{50}{2}\right] + 1$$

$$= \left[1-0{,}94184\right] \times 475 + 1$$

$$= 0{,}05815 \times 475 + 1 = 28{,}6 = \mathbf{29}$$