Retour au site web

Exercice 14-5 : Une illustration de la méthode du maximum de vraisemblance

Appelons f le nombre de familles (12 ici), m le nombre d'enfants par famille (4 dans le texte). Un enfant d'une famille est aléatoirement un garçon ou une fille. La probabilité inconnue d'avoir une fille est p. Le nombre total de filles dans une famille est donc une variable aléatoire X décrite par une loi binomiale

$$p(x) = C_m^x p^x (1-p)^{m-x}.$$

On dispose de $m = 12$ réalisations de la variable X. La vraisemblance de ce résultat est le produit des probabilités d'apparition des valeurs de X observées :

$$V = \prod_1^f C_m^{x_i} p^{x_i} q^{m-x_i}.$$

La « meilleure » valeur de p est celle qui rend maximale la vraisemblance ou, plus commodément, son logarithme.

$$\ln V = \sum_1^f \ln C_m^{x_i} + \ln p \sum_1^f x_i + \ln(1-p) \sum_1^f (m - x_i).$$

Écrivons que la dérivée par rapport à p est nulle :

$$\frac{\partial \ln V}{\partial p} = \frac{1}{p} \sum_1^f x_i - \frac{1}{1-p} \sum_1^f (m - x_i) = 0,$$

soit, en remarquant que $\sum_1^f m = fm$ et en réduisant au même dénominateur :

$$\frac{\partial \ln V}{\partial p} = \frac{1}{p(1-p)} \left(\sum_1^f x_i - pfm \right) = 0.$$

La meilleure valeur de p est donc

$$p^* = \frac{1}{fm} \sum_1^f x_i,$$

soit $p^* = 20/48 \simeq 0{,}42$.