

Chapitre 03 - Biodiversité et son évolution

I - Estimation de l'abondance d'une population

Définition

La **proportion** de certains individus possédant un certain caractère dans une population est égale au quotient de l'effectif de ce groupe d'individu par l'effectif total de la population.

Définition

Un **échantillon** est une partie de la population.

Exemple

Il y a 5 élèves blonds dans une classe de 35 élèves. La proportion d'élèves blonds dans cette classe est

$$\frac{5}{35} = \frac{1}{7} \approx 0,14$$

Propriété

Le dénombrement du nombre d'individus dans un échantillon permet d'estimer l'abondance d'une certaine population, comme les plantes. Mais cette méthode n'est pas adaptée pour les espèces mobiles.

Méthode de capture-marquage-recapture

On cherche à connaître l'effectif N d'une population.

1ère étape: capture-marquage

On capture M individus de cette population. On les marque.

La proportion de ces individus dans la population est $p = \frac{M}{N}$.

2ème étape: recapture

On capture maintenant un échantillon de taille n .

On compte le nombre d'individus marqués: l'effectif compté est m .

La proportion d'individus marqués dans cet échantillon est $p' = \frac{m}{n}$.

3ème étape: conclusion

En supposant que la population ne se modifie pas entre deux captures, la proportion d'individus reste la même: $p = p'$.

Ainsi $\frac{M}{N} = \frac{m}{n}$ ou encore $N = \frac{n \times M}{m}$.

Exemple

On cherche à évaluer l'effectif d'une population de papillons Azuré du Serpolet dans un milieu. Pour cela, 346 papillons sont capturés, marqués puis relâchés. Quelques jours plus tard, on capture 380 papillons l'un

après l'autre que l'on relâche immédiatement. On observe 208 papillons marqués.

Estimer le nombre de papillons Azuré du Serpolet dans le milieu étudié.

Soit N le nombre d'individus dans la population à estimer.

On a:

$$\frac{346}{N} = \frac{208}{380}$$

$$\text{Soit encore: } N = \frac{380 \times 346}{208} \approx 632$$

Il y a environ 632 papillons Azurés du Serpolet dans ce milieu.

II - Estimer la proportion d'un caractère

Le but de ce paragraphe est d'estimer la proportion p d'un caractère dans une population, ne connaissant pas la taille de la population.

définition

La fréquence de m individus dans un échantillon de taille n est $f = \frac{m}{n}$.

Définition

L'ensemble des fréquences de toutes les modalités du caractère s'appelle la **distribution des fréquences** de la série statistique.

Définition

Les distributions des fréquences varient d'un échantillon à l'autre pour une observation sur le même caractère : c'est ce qu'on appelle la **fluctuation d'échantillonnage**.

Même pour des échantillons de même taille, les fréquences peuvent fluctuer.

Propriété

Lorsque la taille n de l'échantillon augmente, l'ampleur des fluctuations des distributions de fréquences calculées sur ces échantillons de taille n diminue et les fréquences tendent à se stabiliser vers des nombres appelés probabilité des modalités du caractère.

Définition

Soit f la fréquence observée d'un groupe d'individus dans un échantillon. L'intervalle de confiance de la proportion p au niveau de 95% est l'intervalle $\left[f - 1,96 \frac{\sqrt{f(1-f)}}{\sqrt{n}}; f + 1,96 \frac{\sqrt{f(1-f)}}{\sqrt{n}} \right]$.

Cet intervalle indique que dans 95% des cas, la proportion p se trouve dans l'intervalle de confiance.

Les étapes de la méthode

1ère étape

On choisit un échantillon de taille n et on calcule la fréquence du caractère observé.

2ème étape

On calcule l'intervalle de confiance au niveau de 95%.

3ème étape

Dans 95% des cas, la valeur p sera contenue dans l'intervalle de confiance calculé à l'étape 2.

Exemple

Une épidémie due à un virus a entraîné la mort d'environ 1000 cerfs de Virginie dans le New Jersey (Etats-Unis). Dans un échantillon de 146 victimes, on dénombre 41 mâles. On veut estimer le pourcentage de cerfs mâles parmi les victimes.

1. Quel est l'intervalle de confiance à 95% du pourcentage de cerfs mâles morts lors de cette épidémie?
On donnera des valeurs approchées à 10^{-3} près par défaut.
2. Conclure.

$$f_{\text{observé}} = \frac{41}{146} \approx 0,325$$

$$I = \left[0,325 - 1,96 \times \frac{\sqrt{0,325(1-0,325)}}{\sqrt{1000}}; 0,325 + 1,96 \times \frac{\sqrt{0,325(1-0,325)}}{\sqrt{1000}} \right] = [0,295; 0,354]$$

On peut dire, avec un risque d'erreur de 5%, que le pourcentage de mâles parmi les victimes est compris entre 29,5% et 35,4%.

III - Modèle Hardy-Weinberg

Le modèle de Hardy-Weinberg permet d'estimer les fréquences alléliques et génotypiques des générations futures pour un gène à deux allèles dans une population.

Ce modèle s'appuie sur les hypothèses suivantes:

- une grande population
- la panmixie (reproduction aléatoire d'individus)
- l'absence de migration, de sélection naturelle et de dérive génétique

On suppose donc qu'un gène possède deux allèles: A (dominant) et a (récessif).

Dans une population, on appelle p la fréquence de l'allèle A et q la fréquence de l'allèle a . On a donc $p + q = 1$ ou encore $q = 1 - p$.

Comme la population est grande, on peut appliquer la loi des grands nombres qui dit que la probabilité pour un parent de transmettre un allèle correspond à sa fréquence.

Ainsi, la probabilité d'avoir un fils possédant le gène A/a est $p \times q$ puisque la proportion d'allèle A dans la population est p et la proportion d'allèle a dans la population est q .

On peut ainsi établir le tableau de fréquence suivant:

		Gamète femelle	
		$A (p)$	$a (q)$
Gamète mâle	$A (p)$	$A//A (p^2)$	$A//a (pq)$
	$a (q)$	$A//a (pq)$	$a//a (q^2)$

On peut donc en déduire de ce tableau la fréquence de l'allèle A à la génération 1.

$$f_1(A) = f(A//A) + \frac{1}{2}f(A//a)$$

$$f_1(A) = p^2 + \frac{1}{2} \times 2 \times pq$$

$$f_1(A) = p^2 + p(1 - p)$$

$$f_1(A) = p^2 + p - p^2$$

$$f_1(A) = p$$

La fréquence de l'allèle A à la première génération sera p , comme à la génération des parents.

De même, la fréquence de l'allèle a à la première génération sera q , comme à la génération des parents.

Le modèle de Hardy-Weinberg permet donc de prédire la structure génétique attendue de la population. On peut déterminer les fréquences de l'allèle A et de l'allèle a et donc des homozygotes $A//A$ et $a//a$ et des hétérozygotes $A//a$.

Exemple

Les loups d'une population ont deux couleurs: noire ou grise.

La couleur de la fourrure est contrôlée par un gène possédant deux allèles: A dominant et a récessif.

Le tableau suivant récapitule le recensement de ces loups selon le génotype.

	$A//A$	$A//a$	$a//a$	Total
Nombre de loups	32	168	600	800
Fréquences observées				

1. Compléter le tableau.
2. Déterminer la fréquence de l'allèle A puis celle de l'allèle a .
3. On suppose que cette population est grande, se reproduit aléatoirement, ne migre pas et qu'elle ne possède pas de sélection naturelle ni de dérive génétique. Déterminer la probabilité d'avoir un fils hétérozygote $A//a$.

1. Voici le tableau complété:

	<i>A//A</i>	<i>A//a</i>	<i>a//a</i>	Total
Nombre de loups	32	318	450	800
Fréquences observées	0,04	0,3975	0,5625	1

$$2. f(A) = f(A//A) + \frac{1}{2}f(A//a)$$

$$f(A) = 0,04 + \frac{1}{2} \times 0,3975$$

$$f(A) = 0,23875$$

$$f(a) = 1 - f(A) = 1 - 0,23875 = 0,76125$$

3. Les hypothèses de Hardy-Weinberg sont vérifiées donc les fréquences de la génération 1 sont les suivantes:

$$f_1(A) = 0,23875 \text{ et } f_1(a) = 0,76125$$

Ainsi la probabilité d'avoir un fils hétérozygote est :

$$f(A//a) = 0,23875 \times 0,76125 = 0,1817484$$