FORMATIONS AUX STATISTIQUES : MANUEL DU FORMATEUR DE NIVEAU 1

Janvier 1991

SCÉNARIO 3

Echantillonnage

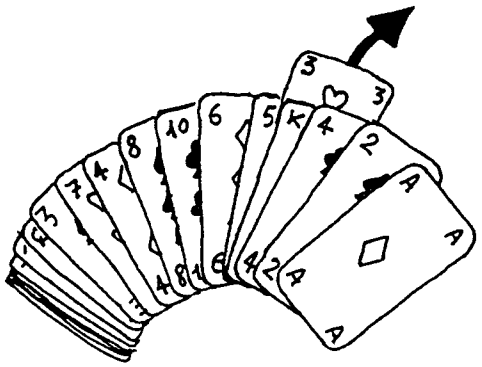
INRA

TRANSPARENTS

DÉPARTEMENT DE BIOMÉTRIE

formation permanente fP

Notion de probabilité



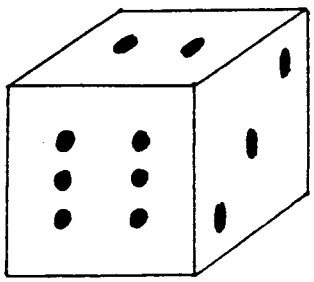
Jeu de 52 cartes :

quatre AS	quatre valets
2	dames
3	rois
.	
.	
10	

Tirer **au sort** 1 carte
au hasard

Quelle **chance** de tirer 1 AS de ♡, 1 ♢?
probabilité

Notion d'équiprobabilité



1 dé - 6 faces
1 lancer de dé

Quelle **probabilité** d'obtenir le 1 ?
Quelle **probabilité** d'obtenir le 6 ?

⇒ Chaque numéro a la **même probabilité** d'être obtenu.

numéro	1	2	3	4	5	6
probabilité	1/6	1/6	1/6	1/6	1/6	1/6

Probabilité = 1 / 6

Notion d'estimation

1 dé - 6 faces

6 lancers

numéro obtenu	1	2	3	4	5	6
fréquence observée	1	1	1	1	1	1

1 expérience : 6 lancers

numéro obtenu	1	2	3	4	5	6
fréquence observée						
probabilité estimée						

$\Downarrow$

estimer la probabilité d'obtenir 1, 2, 3,. . . ,6.

Probabilité estimée $\neq$ probabilité réelle (1 / 6)

Notion de précision d'un estimateur
Notion de précision d'une estimation

Comment améliorer cette situation ?

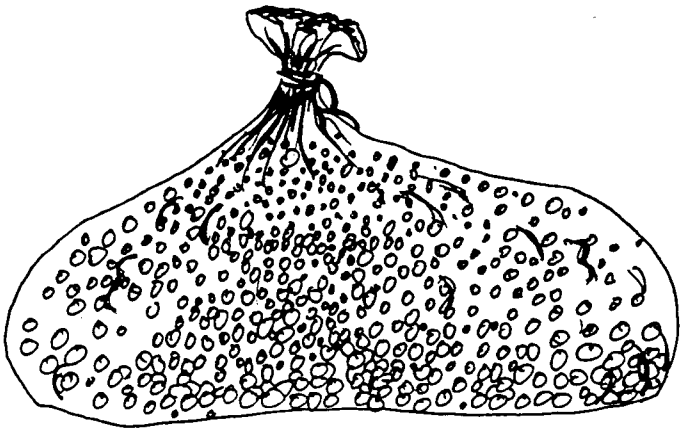
Numéro obtenu	1	2	3	4	5	6
6 lancers						
60 lancers						
60 000 lancers	1/6	1/6	1/6	1/6	1/6	1/6

$\Downarrow$

On "attrape" la probabilité réelle.

6 lancers $\longrightarrow$ 1 estimateur
($\uparrow$ *plus précis*)
60 lancers $\longrightarrow$ 1 estimateur
($\uparrow$ *plus précis*)
60 000 lancers $\longrightarrow$ 1 estimateur

Notion de biais de l'estimation



⟵ grains légers
⟵ grains lourds

Problème : estimer le poids de 1000 grains

Tirer 1000 grains

en haut
grains légers

au fond
grains lourds

Estimation biaisée

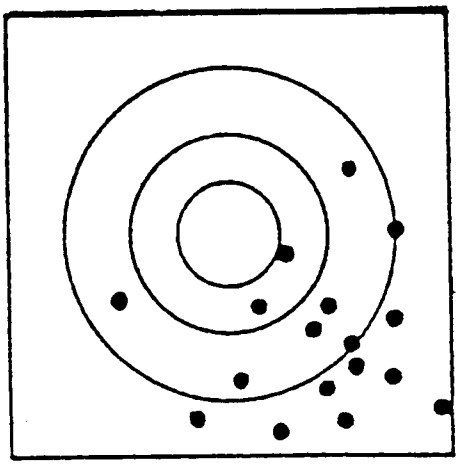
Pourquoi ?

Les grains n'ont pas la même probabilité d'être tirés.

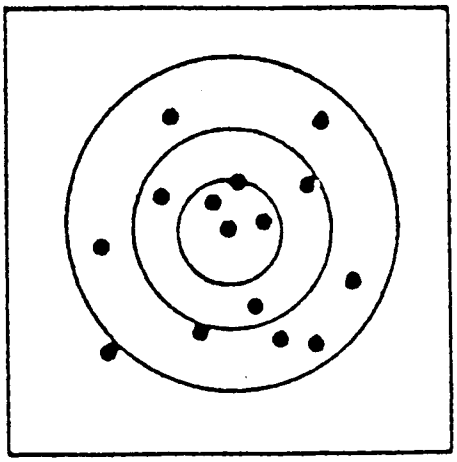
en haut :
plus forte probabilité

en bas :
plus forte probabilité

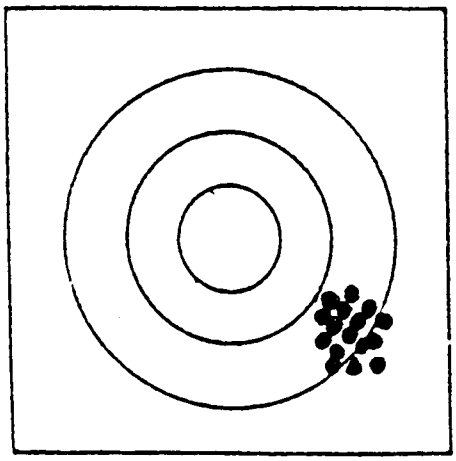
Notion de précision de l'estimateur
Notion de biais de l'estimateur



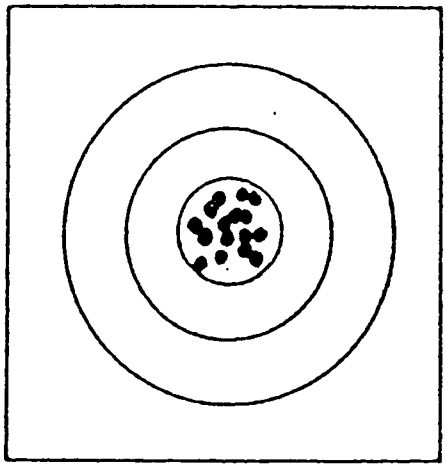
1o) estimateur :
précis - peu précis
biaisé - sans biais



2o) estimateur :
précis - peu précis
biaisé - sans biais



3o) estimateur :
précis - peu précis
biaisé - sans biais



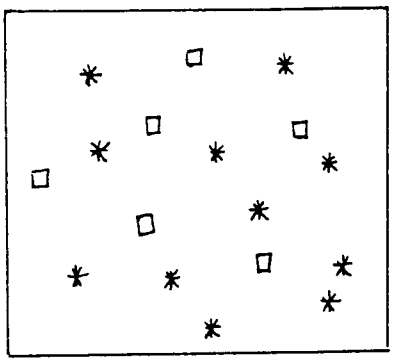
4o) estimateur :
précis - peu précis
biaisé - sans biais

Population = ensemble d'unités statistiques

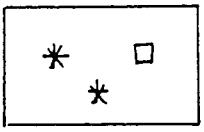
Procedure d'échantillonnage :

- extraire 1 échantillon de la population
- observer certaines caractéristiques

$\Longrightarrow$ en déduire des valeurs concernant l'ensemble de la population



Population

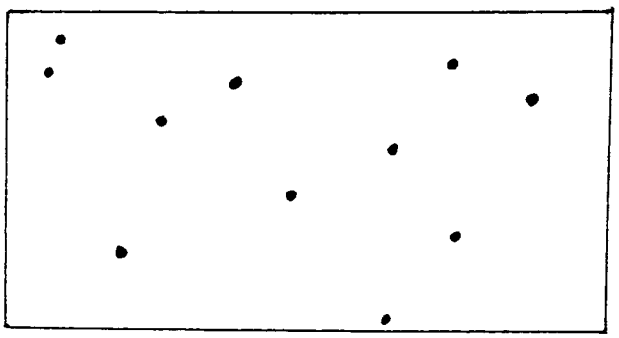


Echantillon

Echantillonnage aléatoire simple

Prélèvement d'unités statistiques :

- **au hasard**
- avec la **même probabilité**
- de façon **indépendante**



1 champ

exemple d'échantillonnage aléatoire simple

Avantages :

- conception simple
- estimateurs no biaisés

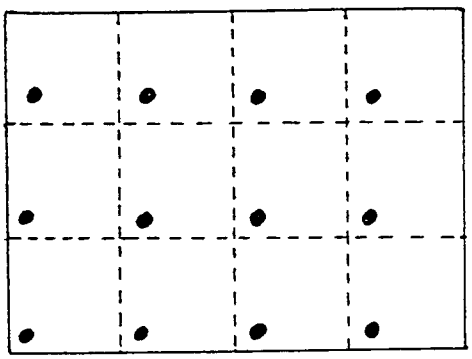
Inconvénients :

- nécessité d'identifier chaque élément de la population
- protocole de sélection des unités pas toujours aisé
- précision faible

Echantillonnage systématique

- n : taille de l'échantillon
- population : n sous-ensembles de même taille t
- tirer au hasard i, entier, entre 1 et t
- prendre la i-ème unité de chaque sous-ensemble

$\Longrightarrow$ échantillon



n = 12

Avantages :

- commodité
- approprié à l'étude de répartition spatial
- évite de prendre des individus proches

Inconvénients :

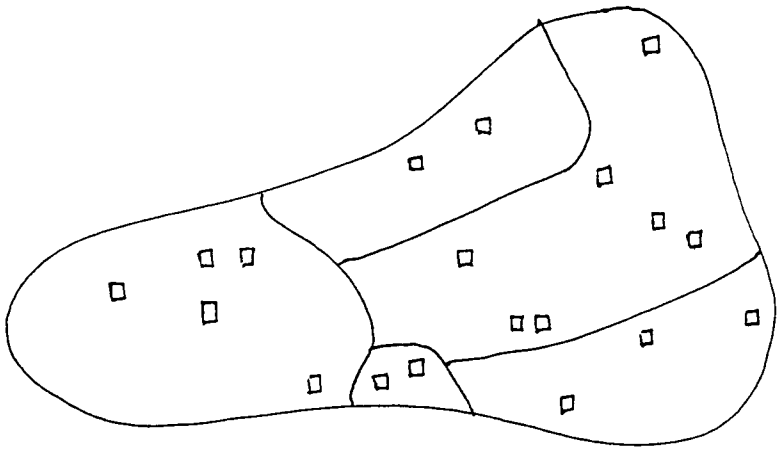
- il faut découper la population en sous-ensembles de même taille
- problèmes si variations périodique

Echantillonnage aléatoire stratifié

région à étudier $\Longrightarrow$ domaines homogènes

population $\Longrightarrow$ strates

Dans chaque strate, échantillonnage aléatoire simple.



Avantages :

- gain en précision,
  surtout si :
  - variabilité inter-strate forte
  - taille d'échantillon dans chaque strate bien choisie.

Inconvénients :

- l'information pour la stratification peut être inadéquate.

Si on ne sait rien

$\Longrightarrow$ **échantillonnage aléatoire simple**
estimateur sans biais,
mais de faible précision

$\Diamond$

Si étude de répartition spatiale

$\Longrightarrow$ **échantillonnage systématique**

$\Diamond$

Si on a une connaissance de la population,
Si on peut la diviser en strates homogènes

$\Longrightarrow$ **échantillonnage aléatoire stratifié**
(plus précis que l'échantillonnage aléatoire simple)

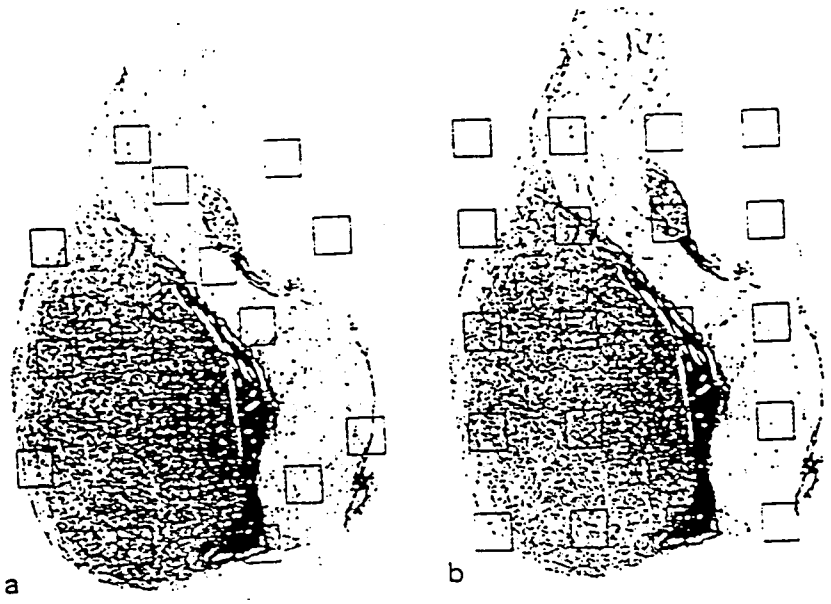


fig.1 : microscopie optique : coupe d'une hypophyse

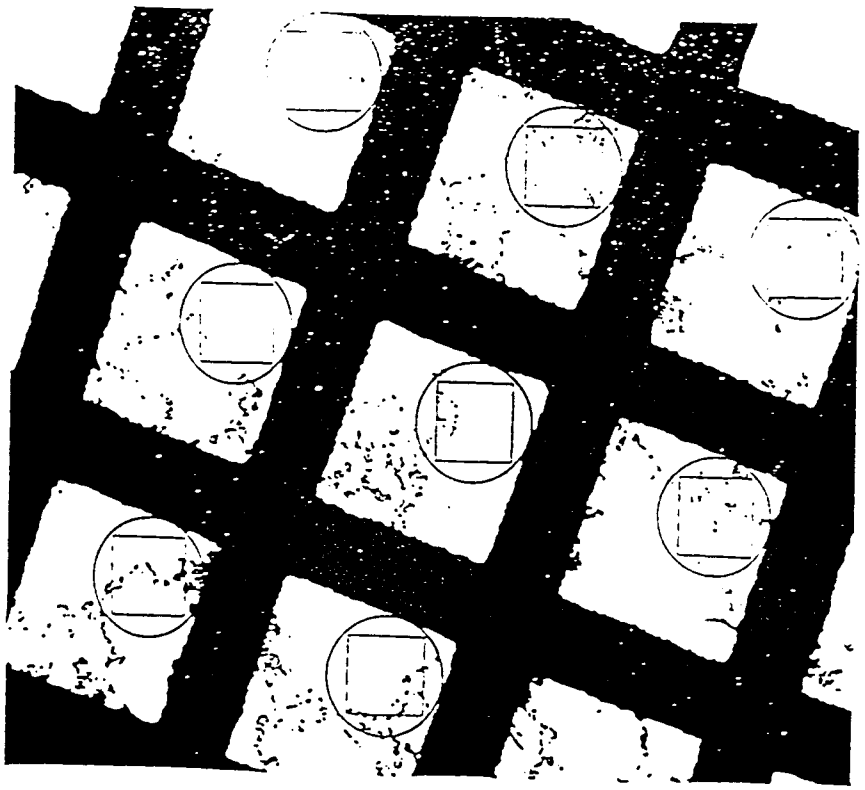


fig.2 : microscopie électronique : échantillonnage systématique

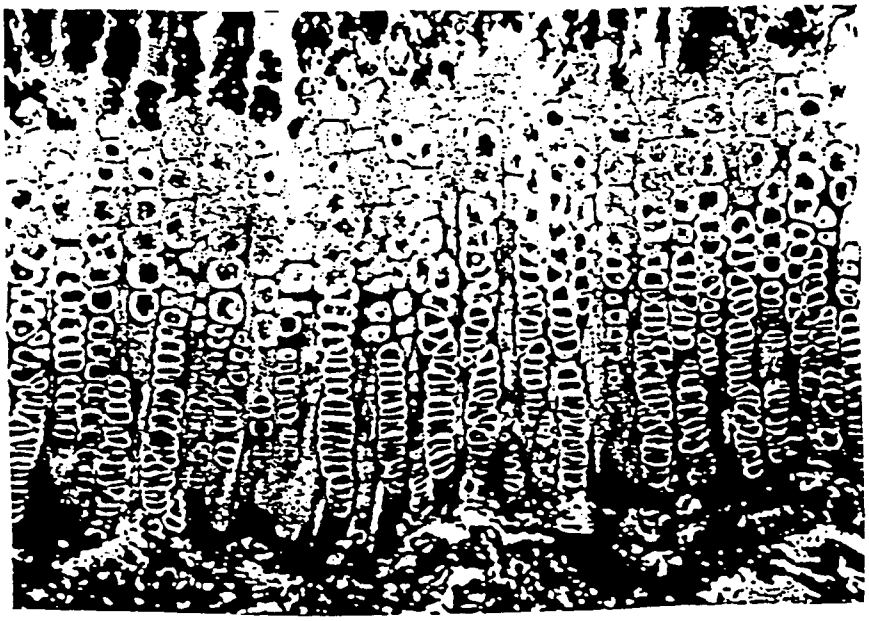


fig.3 : coupe dans le cartilage d'un os

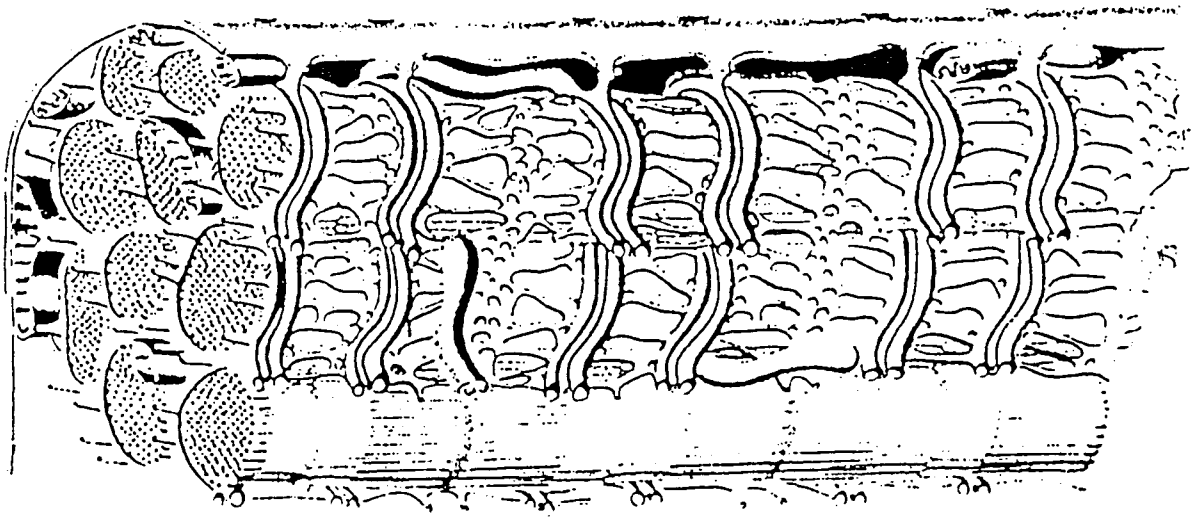


fig.4 : muscle du squelette

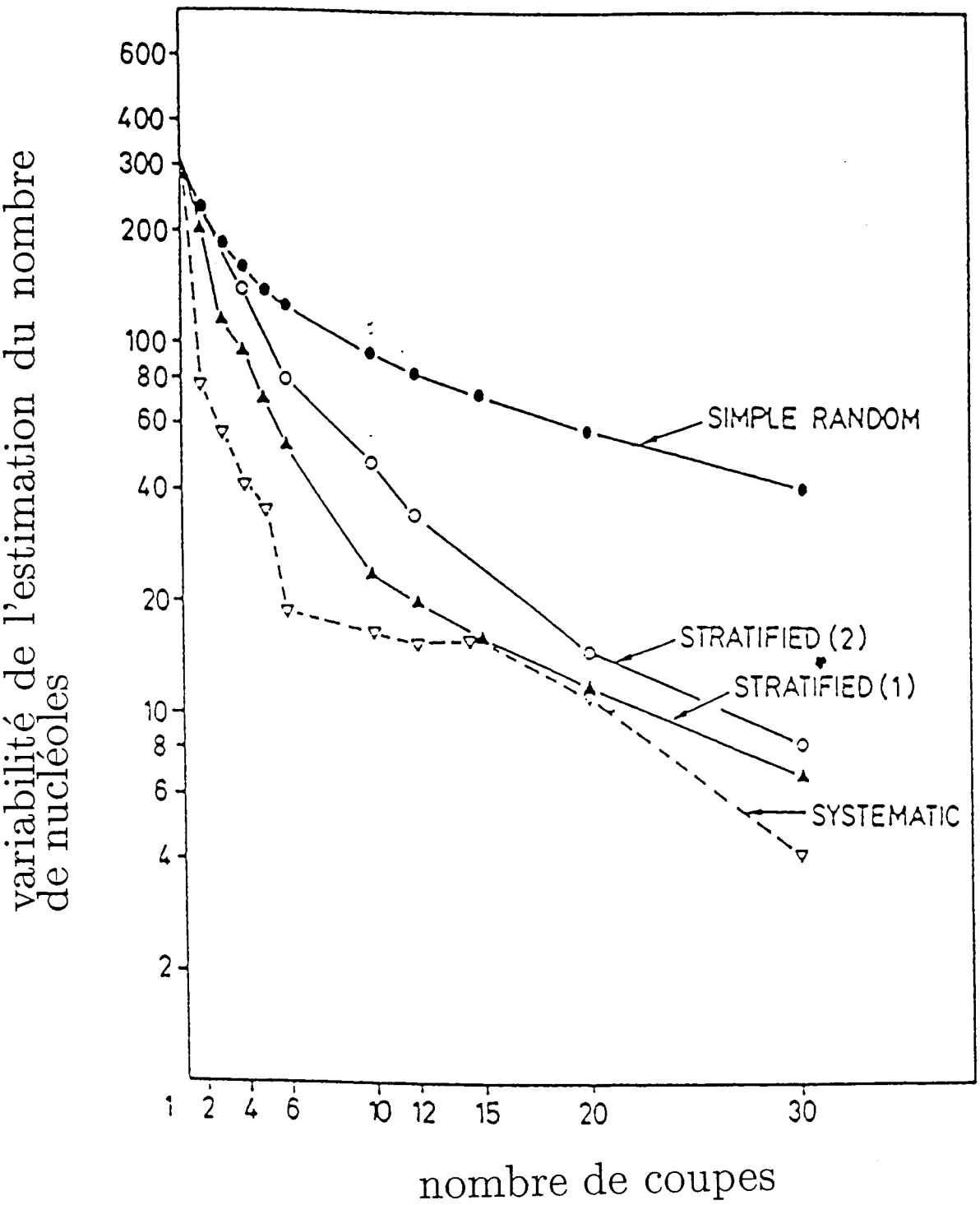


fig.5 : comparaison de la précision de quatre types d'échantillonnage

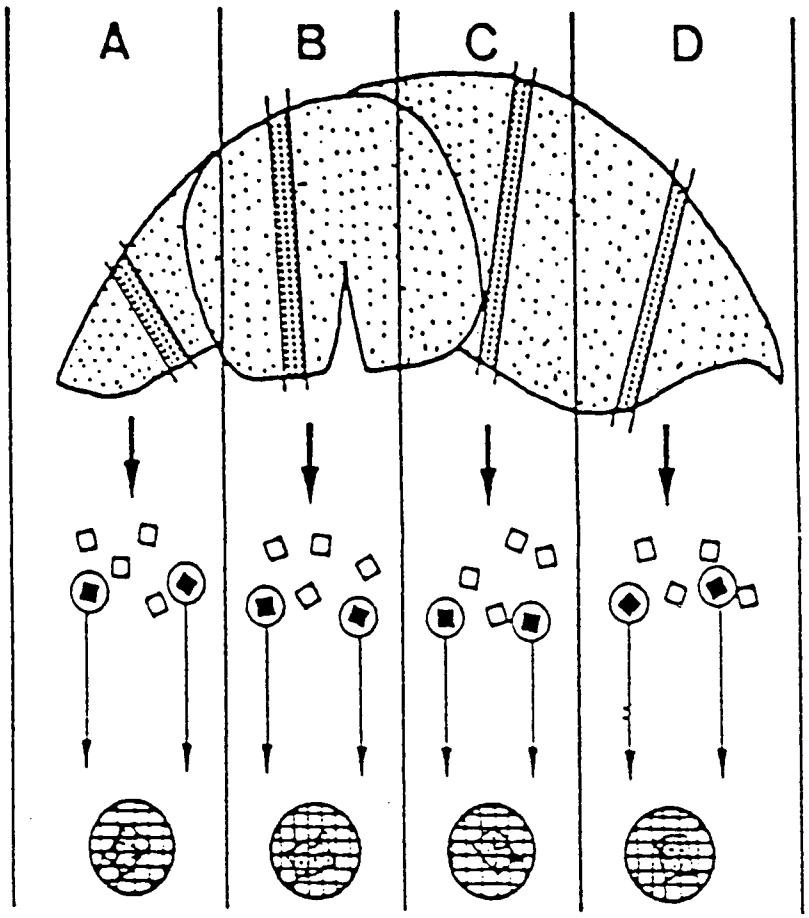
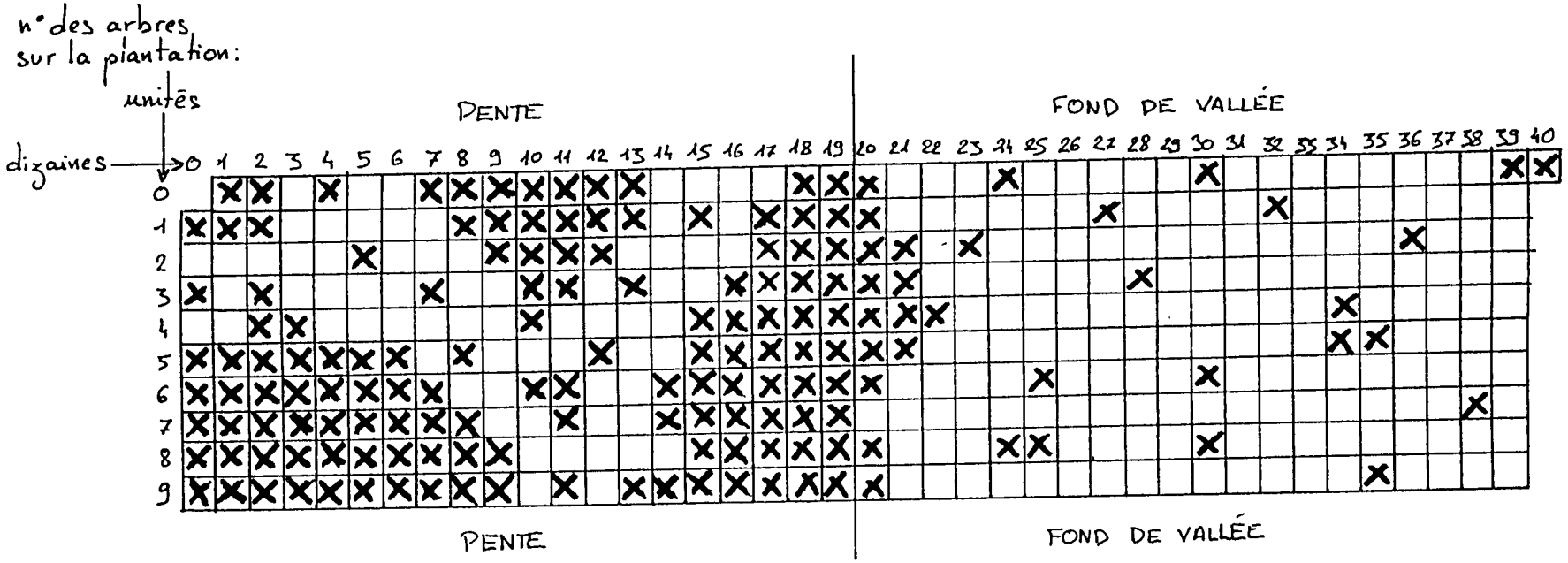


Fig. 6 : un organe, schématisation du mode d'échantillonnage

plan du verger



Légende :

☐ arbre faiblement infesté

☒ arbre fortement infesté

ZONE DE PENTE :

71 arbres faiblement infestés (note de 1 à 3)
128 arbres fortement infestés (note de 4 à 5)

ZONE DE FOND DE VALLÉ :

168 arbres faiblement infestés (note de 1 à 3)
33 arbres fortement infestés (note de 4 à 5)

ENSEMBLE DU VERGER :

239 arbres faiblement infestés (note de 1 à 3)
161 arbres fortement infestés (note de 4 à 5)