Mesure et échantillonnage

Introduction à l'approche stéréologique

Hugues Talbot

talboth@esiee.fr

ISBS / ESIEE 2005

Origine de ce document

- Texte d'une conférence invitée de Adrian Baddeley, professeur de mathématique à l'UWA, Perth, la "Bahadur memorial lecture", en l'honneur de Raj Bahadur, professeur de statistiques à l'université de Chicago.
- A. Baddeley a été un moment VP de la société internationale de stéréologie pour l'Australasie.
- Je me suis contenté de traduire, reprendre les exemples et en compléter certains.
- Ceci est le premier cours donné aux élèves ISBS 3ème année en spécialité bio-imagerie. Je ne présente aucun résultat de recherche récent.

Chaîne de l'analyse d'image

1. Acquisition d'image – cf cours P-XII ;
2. Filtrage – cf cours morpho 2e année + TDSI cette année ;
3. Segmentation – cf cours morpho 2e année + TPs/Projets cette année ;
4. Mesure – ce cours.

Pb de la mesure en général

- En général on cherche à obtenir de l'information, par exemple un nombre d'objets, une surface, etc.
- On veut que cette mesure corresponde à qqchose de physique ;
- On veut éviter les erreurs. Il y a deux types d'erreurs
  1. Les erreurs aléatoires (dues à l'instrumentation, au bruit, à la discretisation, etc).
  2. Les erreurs systématiques. les *biais*.

Comment éviter les erreurs ?

- Eviter les erreurs aléatoires :
  - Multiplier les mesures
  - Filtrer (moyenne, médian, etc)
- Eviter les erreurs systématiques :
  - Faire attention au protocole
  - Réflechir !
  - Appliquer de bonnes méthodologies, en particulier la *stéréologie*.
  - En particulier, une bonne approche stéréologique commence par l'échantillonnage.

Principes d'échantillonnage

La stéréologie, ou savoir compter

Une histoire du Mahabharata (1000 ans AVJC).

Merci à A. Baddeley, UWA, dpt of mathematics.

L'histoire de Nala

Nala était un roi qui ne songeait qu'au jeu.
Il perdit tout sauf sa femme fidèle et ses talents de cavalier.
Humilié, il se rendit au service du bon roi Rtuparna.
Rtuparna admirait les talents de cavalier de Nala, et voulu l'apprendre de lui.
Pour inciter Nala à lui révéler ses secrets, Rtuparna décida d'impressionner Nala avec ses talents à lui : compter.

Nala (II)

Ils galopèrent en passant devant un arbre Vibhitaka.

Rtuparna prit quelques brindilles de l'arbre, et déclara :

Cet arbre port 2095 fruits et 100.000 feuilles.

Nala (III)

Nala insista qu'ils s'arrêtent et comptent les fruits. Après un long, long moment, ils arrivèrent à presque le même nombre.
Complètement abasourdi, Nala demanda comment Rtuparna avait fait pour compter si vite.
Rtuparna répondit:

> Je possède la science des dés, et donc je suis doué pour les nombres.

Par la suite, Nala et Rtuparna se mirent d'accord pour échanger leurs secret, mais à voix basse...

Nala (fin)

Ce que fit probablement Rtuparna:

- Prendre un échantillon de quelques branches représentant $\frac{1}{m}$ de l'arbre ;
- compter les fruits sur ces branches ;
- estimer

$$\text{fruits sur l'arbre} = m \times \text{fruits comptés}$$

Principes d'échantillonnage

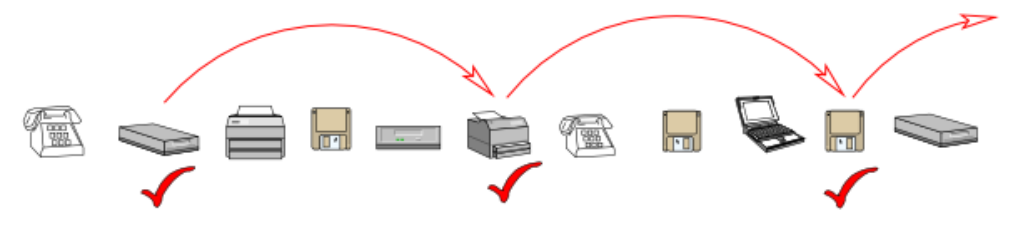
- Soit une population d'unité distinctes :

- On *échantillonne* une sélection d'unités de la population

Échantillonnage systématique

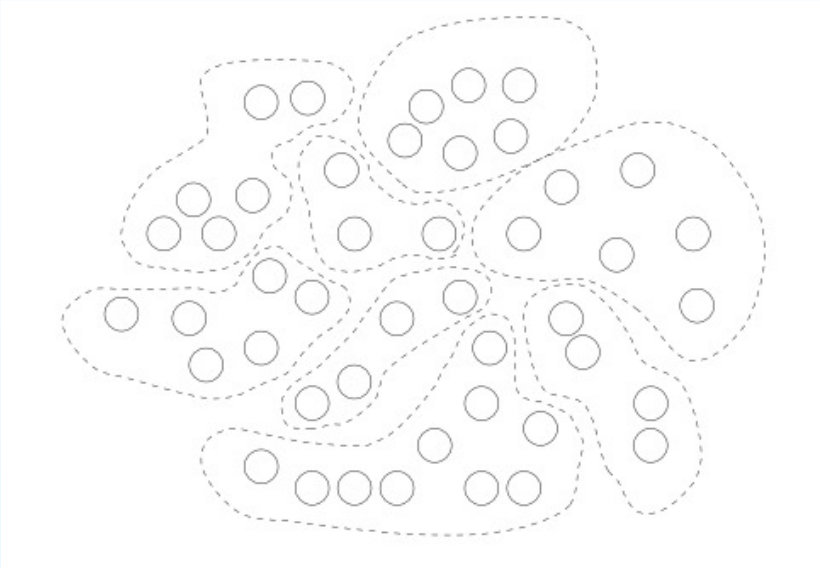
Il est essentiel d'obtenir un échantillon *aléatoire* et *représentatif* de la population totale. Pour obtenir un échantillon $1/m$ de la population totale, on peut:

- Ordonner les unités dans une séquence arbitraire ;
- Choisir un nombre aléatoire entre 1 et m ;
- Utiliser ce nombre comme le point de départ, et prendre un objet tout les m :



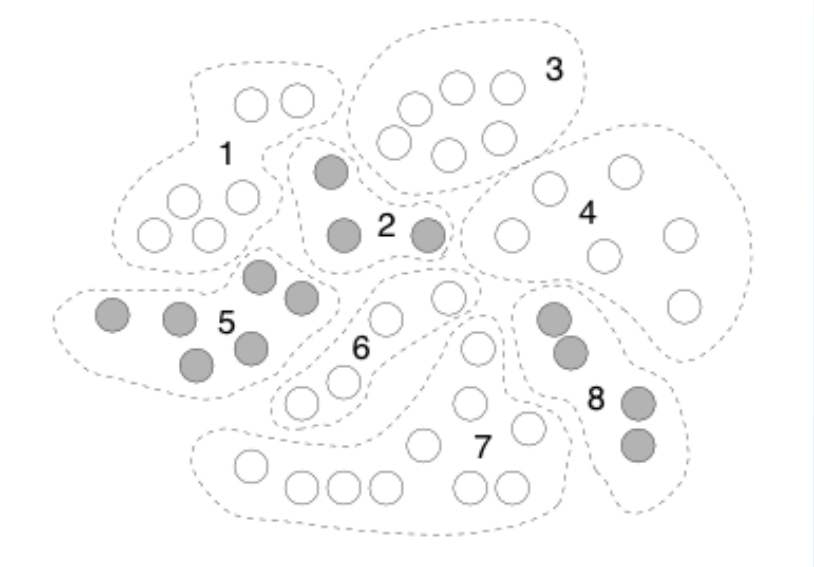
Echantillonnage en Cluster

Si la population s'organise naturellement en groupes (clusters)



Echantillonnage en Cluster

1. Grouper la population en clusters
2. Traiter les clusters comme l'unité d'échantillonnage et effectuer un échantillonnage systématique.
3. Traiter toutes les unités dans les clusters sélectionnés comme les échantillons.

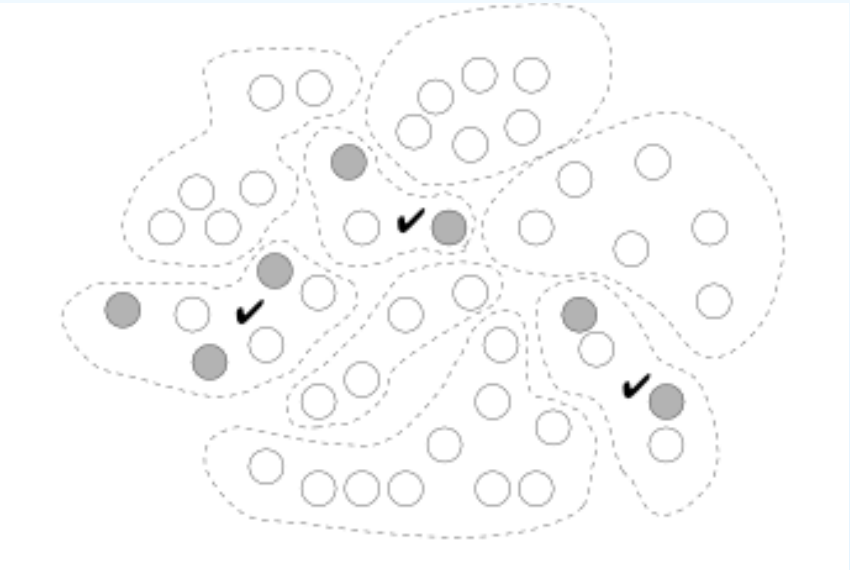


Exemple

Les feuilles d'un arbre se groupent naturellement en branches
On peut prendre un échantillon des feuilles sur l'arbre en prenant un échantillon de branches.

Echantillonnage deux étapes

1. Grouper la population en clusters
2. Traiter les clusters comme l'unité d'échantillonnage et effectuer un échantillonnage systématique.
3. À l'intérieur de chaque cluster, opérer un échantillonnage systématique.



Compter les mots dans un livre

- Prendre tout les 4 chapitres
- Prendre la 10ème page dans chaque chapitre
- Prendre la 5ème ligne sur chaque page
- Compter tous les mots sur ces lignes et les additionner
- Multiplier le tout par 200.

Ce que fit Ruparna?

1. Echantillonner sur les m-ème branches
2. Prendre toutes les k brindilles sur ces branches
3. Compter les fruits sur toutes ces brindilles
4. Multiplier le tout par $m \times k$ pour obtenir le total.

Sauf que...

$$2095 = 5 \times 419$$

et 419 est premier...

Alors?

En imaginant qu'il y ait eu 5 fruits dans l'échantillon de Rtuparna, alors il a pu décider que ces 5 fruits n'étaient plus sur l'arbre après qu'il ait eu enlevé la brindille qui les portait.
Dans ce cas il y aurait eu

$$2095 + 5 = 2100$$

fruits dans l'arbre originellement, et $2100 = 5 \times 420$, et

$$420 = 2^2 \times 3 \times 5 \times 7$$

La fraction d'échantillonnage était donc sans doute $1/420$...

Les cellules du cerveau

Maladies du cerveau

- Les maladies du cerveau chez les personnes âgées (Démence sénile) et autres désordres organiques sont une source majeure de coût de santé publique.
- les années 1990s furent déclarées "décade du cerveau" (avec fonds associés)
- Motivation majeure:

  Le cerveau perd progressivement des neurones avec l'âge
  Seul parmi tous les organes, le cerveau apparaît incapable de régénérer ses propres tissus.

Article

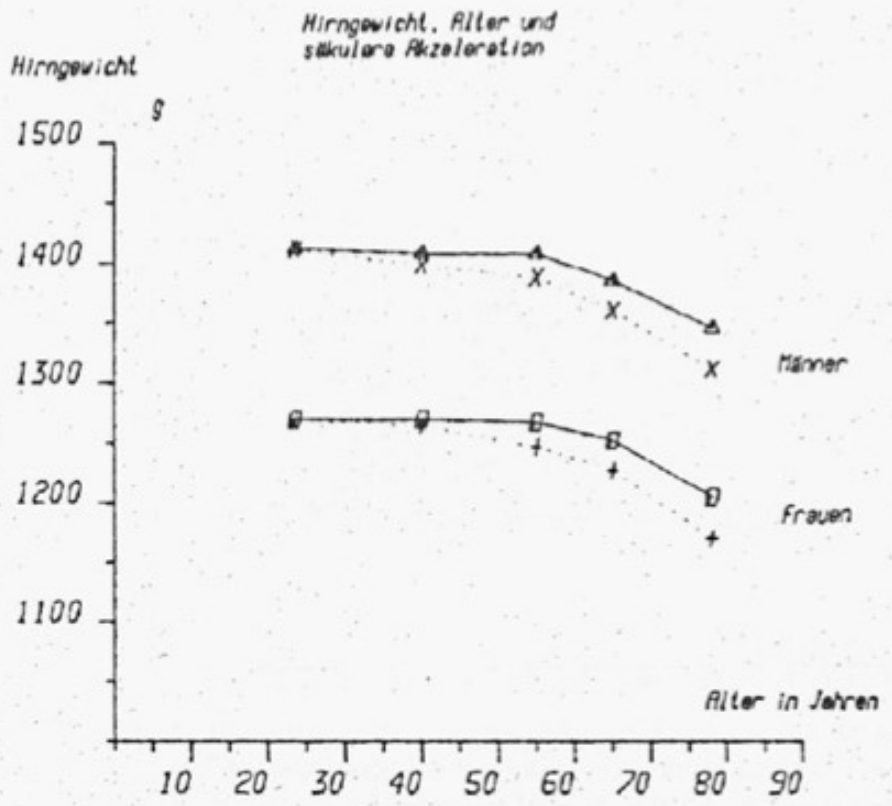


Abb. 1 Das mittlere Hirngewicht steht in der Ordinate und das Alter, wie in den folgenden Abbildungen, in der Abszisse. Die Werte der gestrichelten Linie entsprechen den Mittelwerten von insgesamt 24000 Gehirnwägungen. Die durchgezogene Linie berücksichtigt die säkulare Akzeleration und entspricht der Veränderung des Hirngewichtes einer durchschnittlichen Person während des Lebens.

Etudes

- De nombreuses études ont montré que la densité de neurones (cellules par cm^3) diminue avec l'âge.
- Ces études étaient toutes basée sur des petits échantillons de tissus du cerveau.

Contre-étude

Nervenheilkunde
© F. K. Schattauer Verlag GmbH (1985)

Aus Wissenschaft und Forschung

Gibt es Nervenzellverluste während der Alterung in der menschlichen Hirnrinde?

Ein morphometrischer Beitrag zu dieser Frage*

H. Haug

Aus dem Institut für Anatomie der Medizinischen Hochschule Lübeck (Direktoren: Prof. Dr. med. H. Haug – Prof. Dr. med. W. Kühnel)

Haug 1985, M&M

- Des tissus frais sont préparés pour la microscopie par fixation, colorisation et montage.
- Il est bien connu que les tissus mous rétrécissent à la fixation
- Toutes les mesures précédentes de densité de cellules étaient pour les tissus *traités*.
- On constate que les tissus jeunes rétrécissent plus que les tissus vieux.

Nouvelle estimation après correction

Si on tient compte de cette différence de rétrécissement...

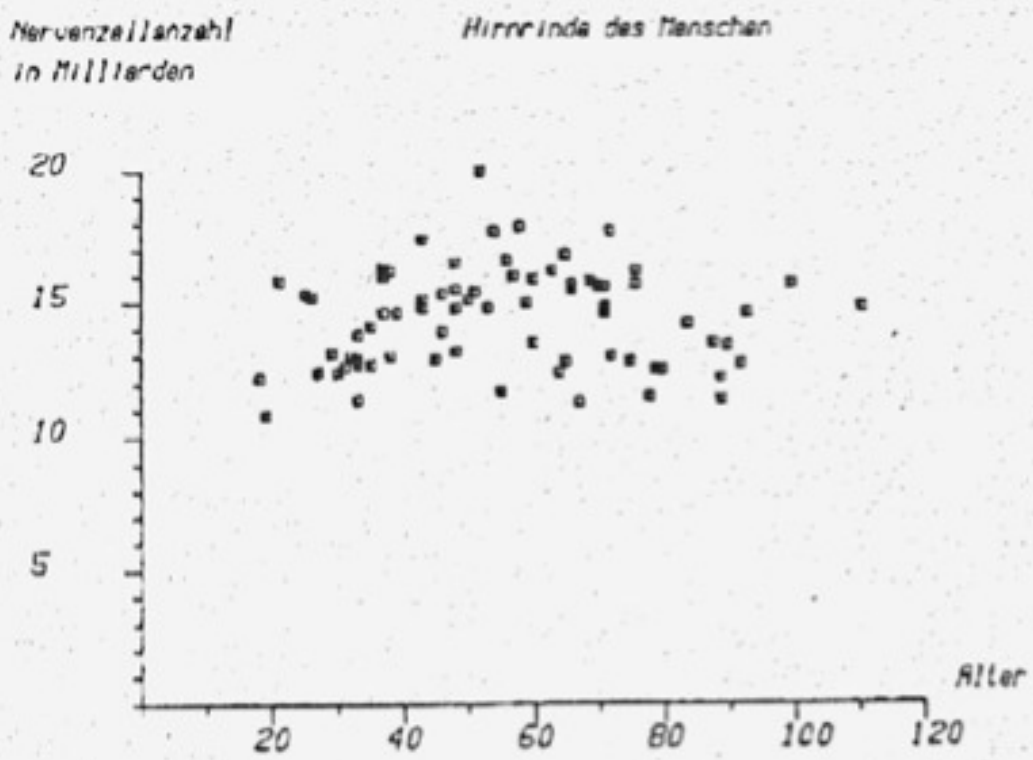


Abb. 5 Die Gesamtzahl aller Neuronen der Großhirnrinde ist in der Ordinate aufgeführt (jeder Wert enthält das Mittel aus mindestens 2 Arealen).. Die Werte wurden aus einer normierten mittleren Zelldichte und der individuellen Hirngröße errechnet. Keinerlei Trend zur Änderung der Anzahl der Neuronen kann während der Alterung festgestellt werden.

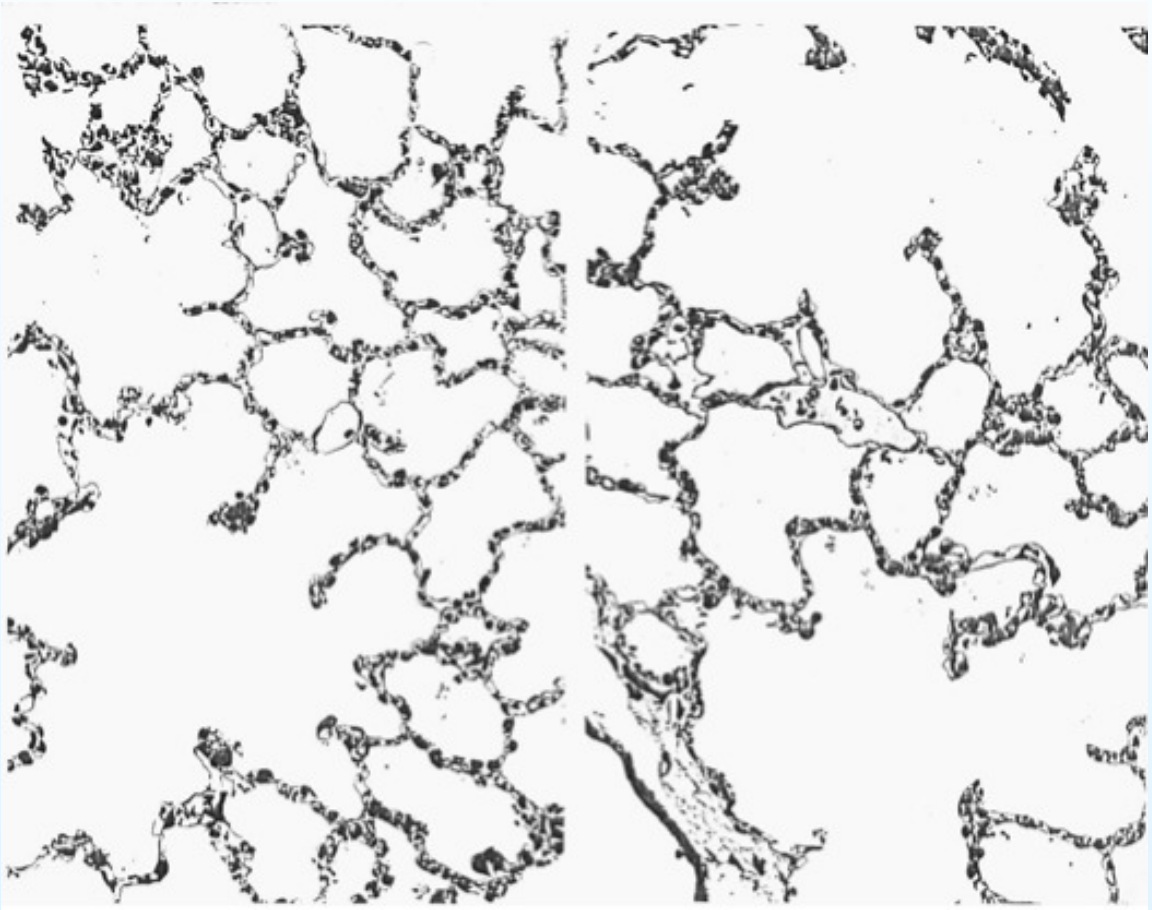
Un biais

Toutes les "preuves" de la perte de neurones avec l'âge se sont révélées être un artefact dû au rétrécissement.
Vers environ 1995, de nouvelles expériences ont démontré la prolifération de neurones dans le cerveau.
Prolifération avait été démontrée auparavant chez le rat.

Stéréologie

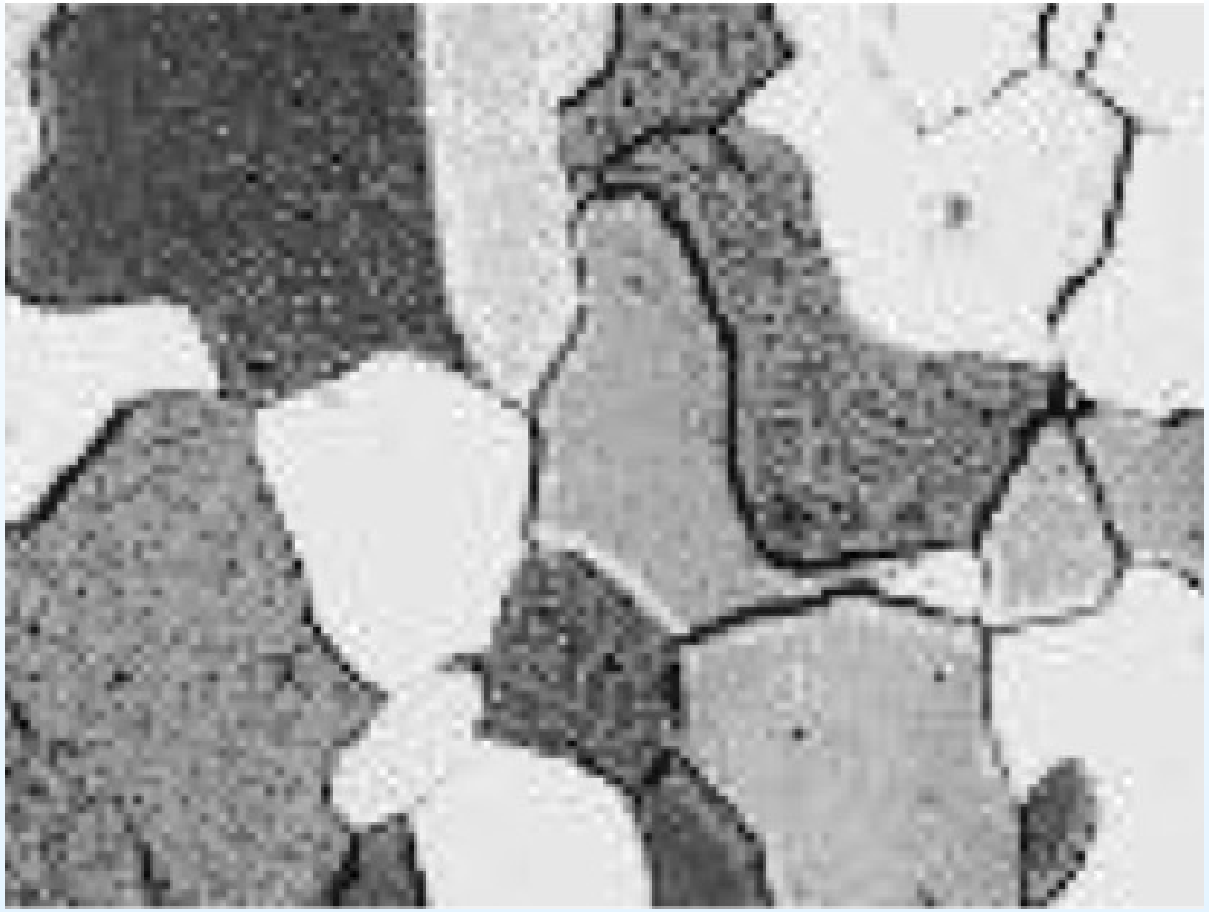
Origine du nom ”stéréologie”

Le nom provient d’une technique d’obtention de données en 3D à partir de coupes microscopiques 2D.



Delesse (1847)

Microscopie de coupes polies de roches et minéraux



Idée: La composition d'une section est representative de la composition de la roche.

Relation de Delesse

Si on souhaite la fraction de quartz par volume:

$$V_v = \frac{\text{volume total de quartz}}{\text{volume total de roche}}$$

On observe la proportion de quartz en surface dans une section

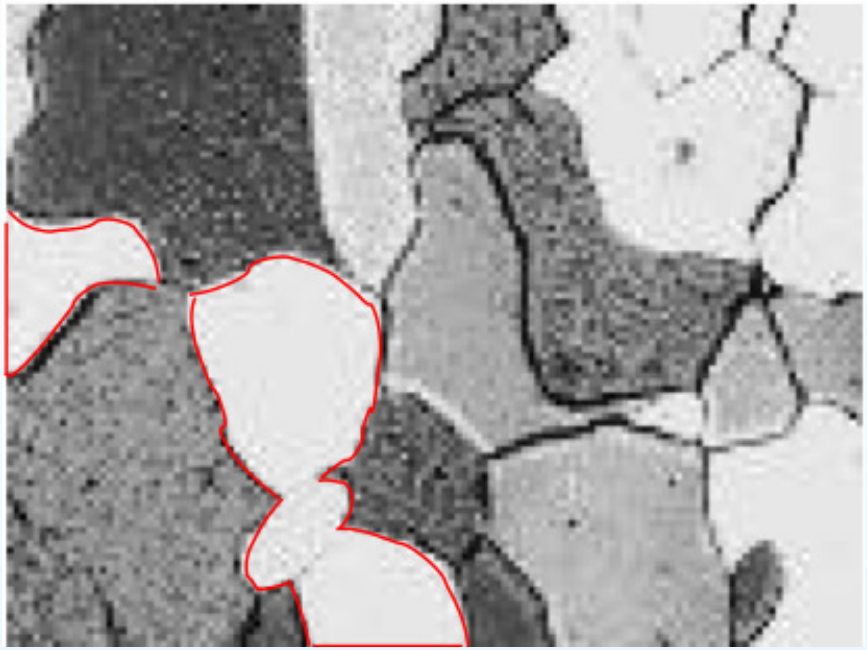
$$A_a = \frac{\text{Aire observée occupée par le quartz}}{\text{Aire de la section}}$$

Delesse écrit:

$$V_v = A_a$$

Exemple

A l'analyseur d'image, on mesure la fraction blanche en surface:

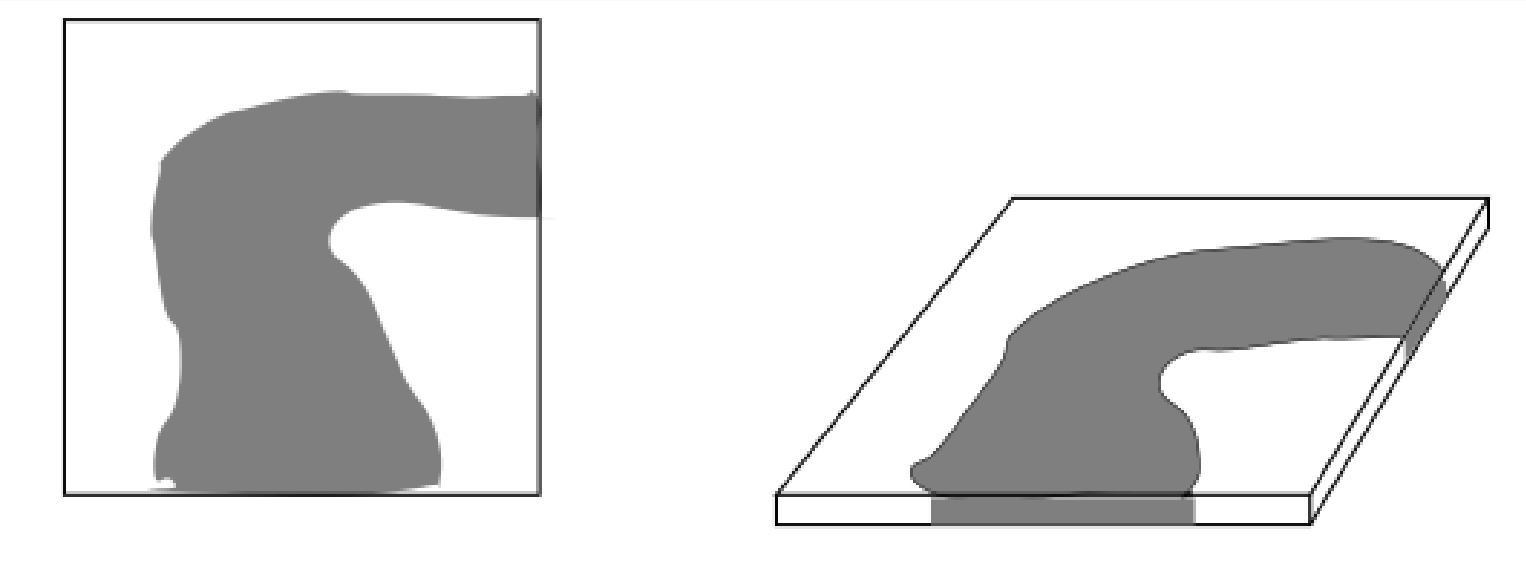


$$A_a = 0.13$$

Par la formule de Delesse, on estime qu'il y a 13% de quartz dans cette roche.

Justification

On admet qu'une section plane est comme une tranche très fine.



Le volume de chaque phase est égal à l'aire relative multipliée par l'épaisseur de la tranche.

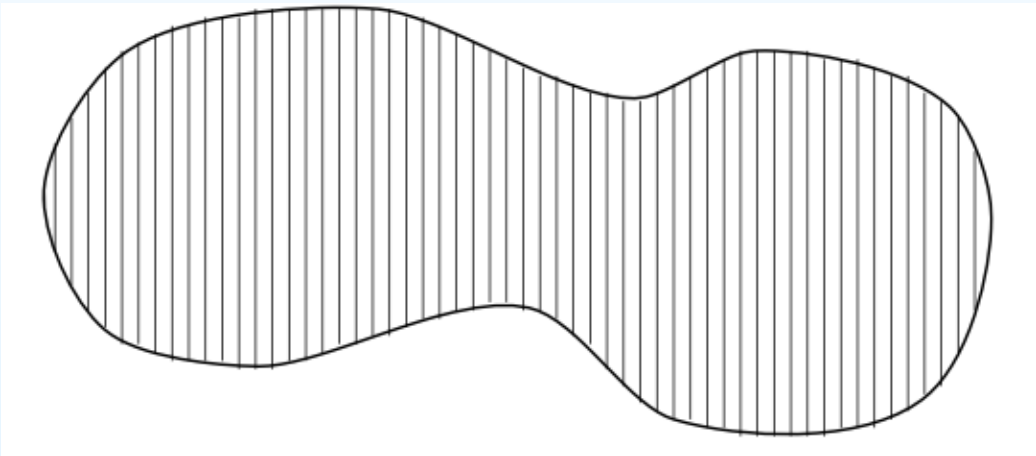
$$V = A \times t$$

Pour les matheux

Théorème de Fubini, pour $X \in I\!R^3$

$$\text{volume}(X) = \int \text{aire}(X \cap T_h)dh$$

Où $T_x = \{(x, y, z) : z = h\}$



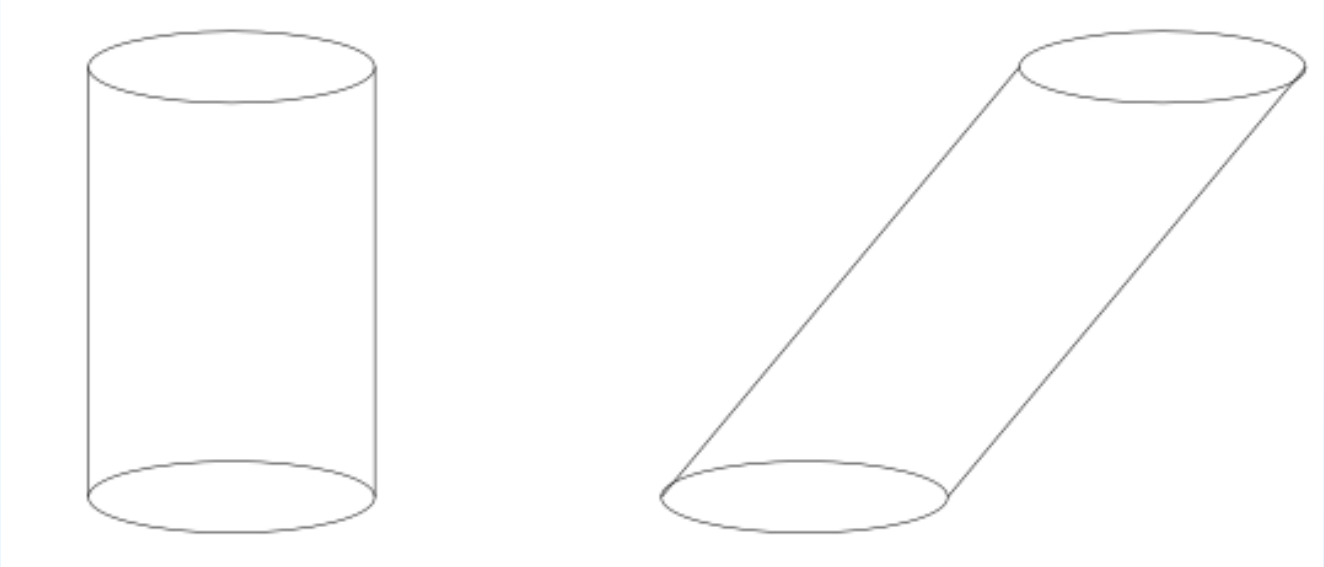
Propriétés

La formule de Delesse $V_v = A_a$:

- Utilise des principes simples de géométrie et échantillonnage
- Ne dépend pas des la forme des objets/phases
- N'est pas spécifique aux matériaux
- Ne requiert pas une reconstruction 3D.

Principe de Cavalieri

Deux solides avec des sections identiques en chaque plan horizontal ont le même volume:



Microscopie quantitative en 1840

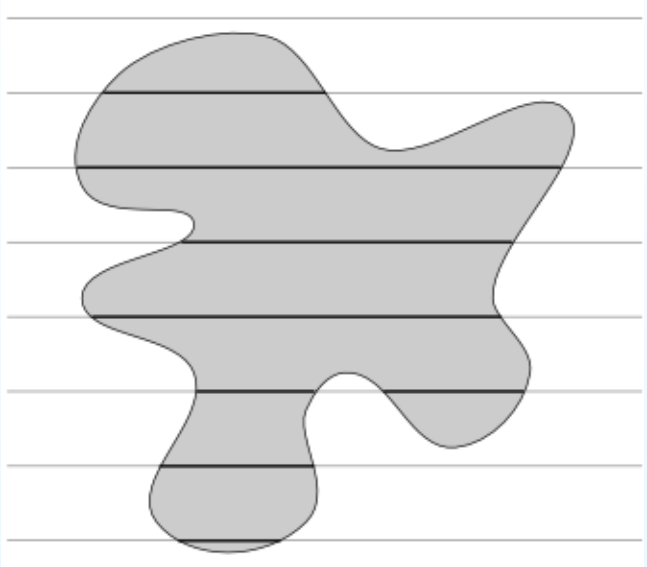
En 1840, la mesure de A_a était faite de la façon suivante:

1. Tracé de la section sur carton
2. Découpage des formes
3. Pesée...

Peut-on imaginer plus pratique ?

Lignes de Rosiwal (1889)

Rosiwal surimpose une grille de lignes de test



et mesure la fraction de longueur:

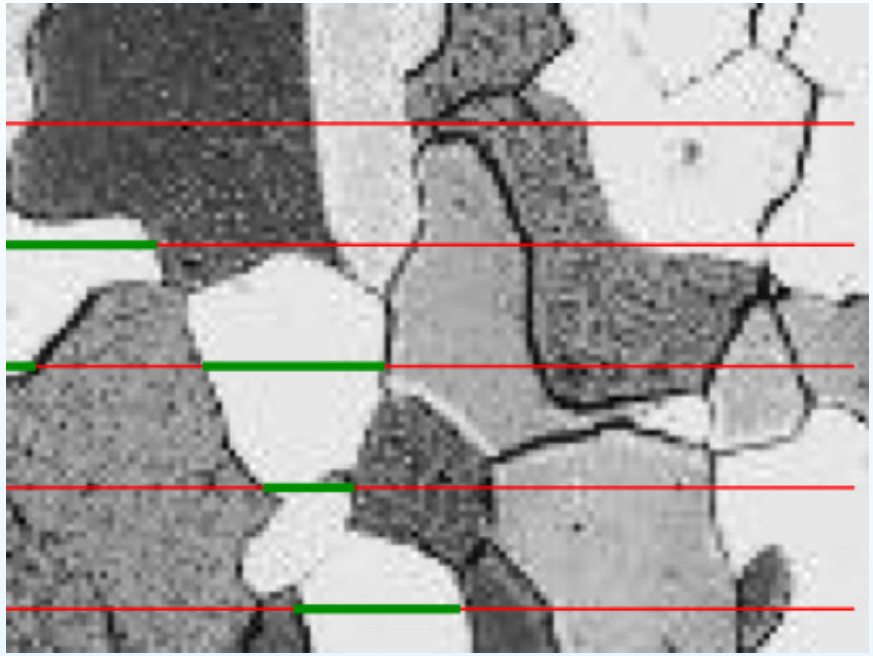
$$L_l = \frac{\text{longueur totale des ligne grisée}}{\text{longueur totale de toutes les lignes}}$$

et ensuite :

$$A_a = L_l$$

Exemple

A l'analyseur d'image, la phase blanche occupe 15% des lignes de test.

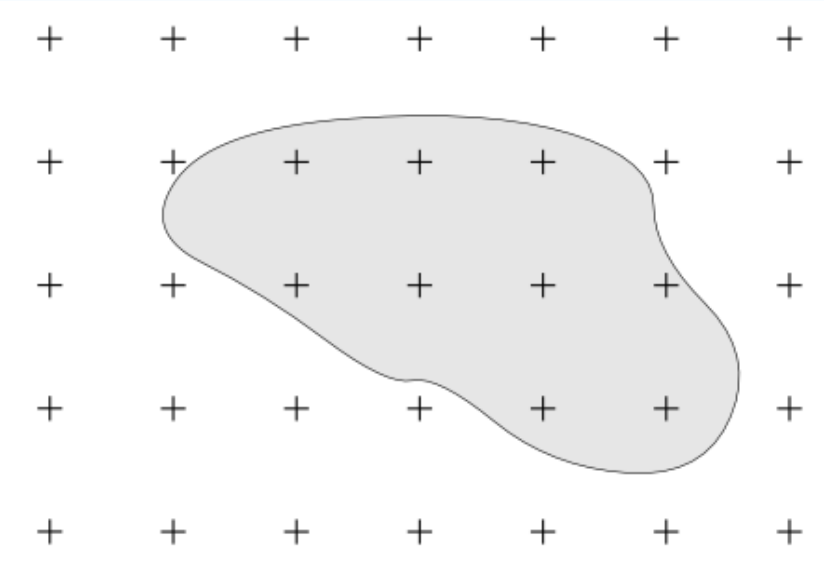


$L_l = 0.15$

On estime $V_v = 0.15$. La phase blanche occupe 15% du volume de roche.

Encore plus simple

Glagoleff (1943) surimpose un réseau de points.



On mesure la fraction de points comptés

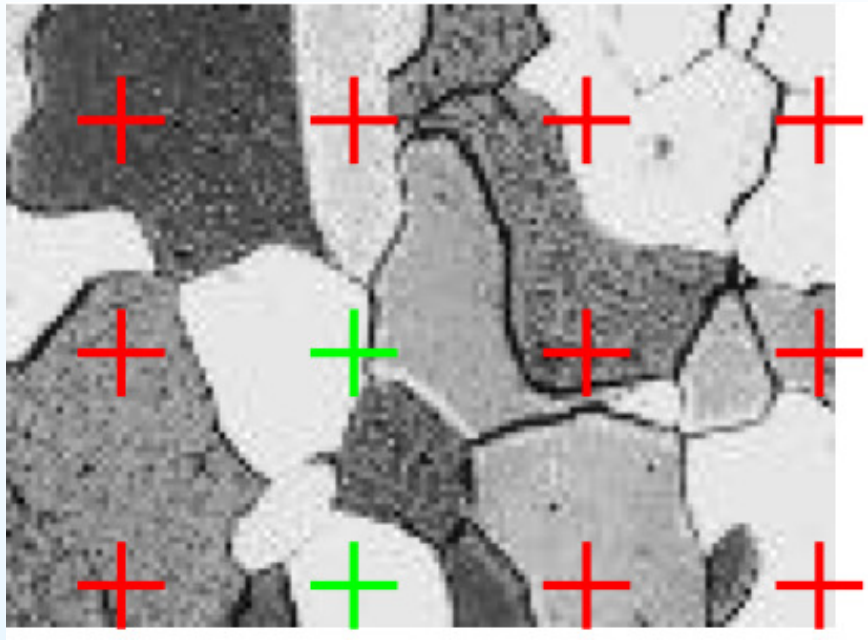
$$P_p = \frac{\text{nombre de points dans la partie grisée}}{\text{nombre de points totaux}}$$

Puis

$$A_a = P_p$$

Exemple

2 points sur 12 tombent dans la phase blanche. $2/12 = 0.17$



$P_p = 0.17$

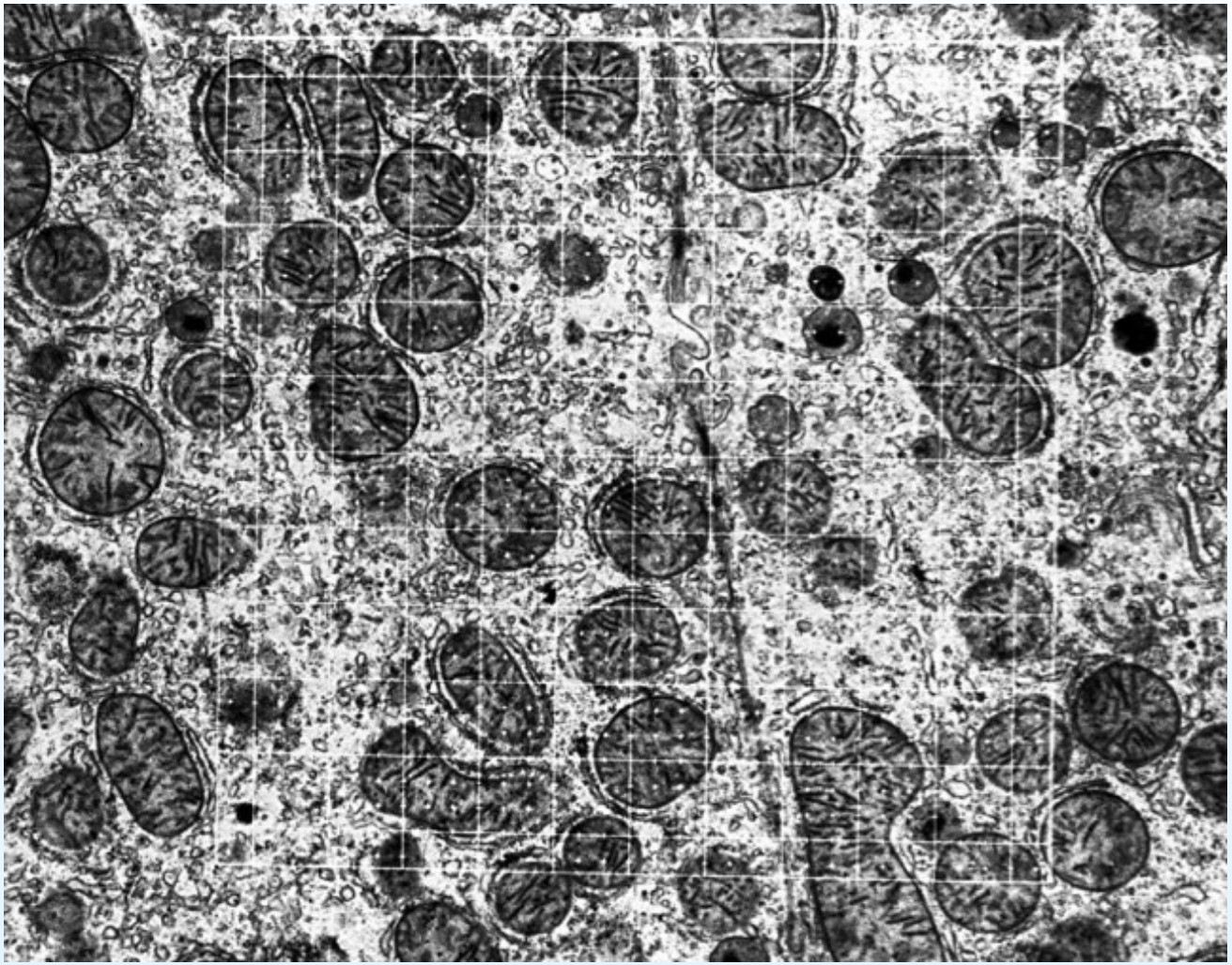
On estime que $V_v = 0.17$.

En résumé

Méthodes stéréologiques pour déterminer la fraction volumique:

$$V_v = A_a = L_l = P_p$$

Système de test



Méthode d'échantillonnage systématique.

Paradoxe de la variance

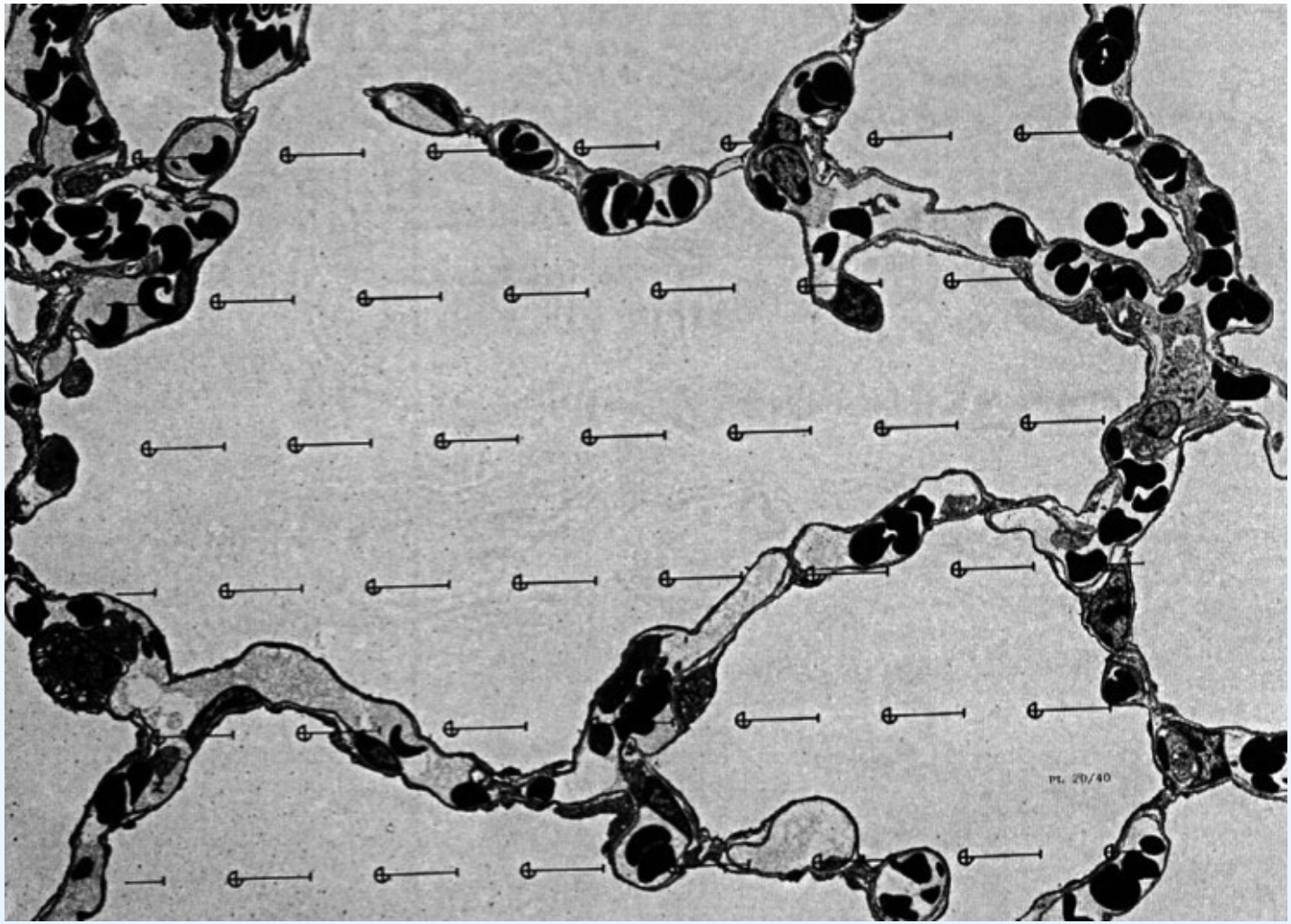
L'utilisation de système de test n'est plus à la mode, on préfère l'analyse d'image.
Mais il arrive que $\text{var}(A_a) > \text{var}(P_p)$ en pratique, du fait de la forte corrélation entre pixels voisins.

Autres systèmes de test

Il existe des techniques stéréologiques pour mesurer

- Les surfaces 3D (membranes, interfaces, frontières)
- Les longueurs (filaments, tubules, dendrites)
- Les épaisseurs (couches, etc)
- Les tailles moyennes (cellules, grains, particules, etc)
- La connectivité d'un réseau, tubes, fibres, etc.

Exemple - connectivité



Stéréologie et comptage

Dénombrement

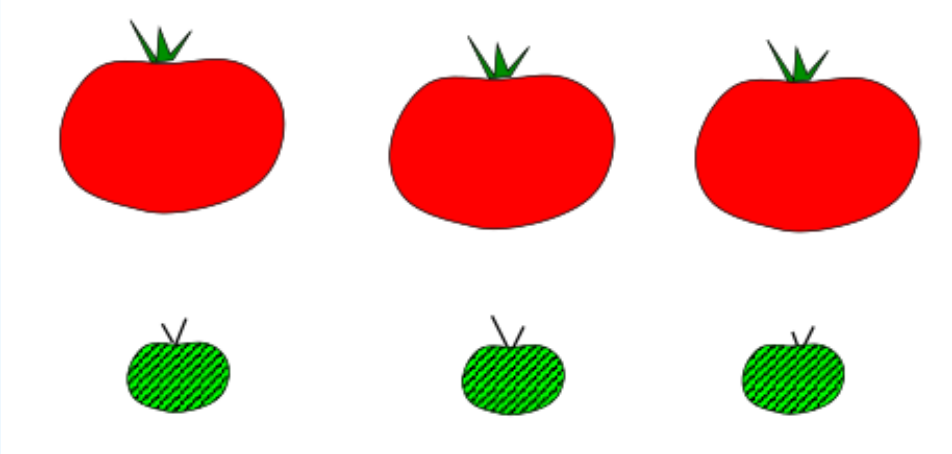
La stéréologie classique n'inclut pas de techniques pour estimer des nombres avec la même généralité.
Erreurs classiques:

- Volume moyen des cellules = 234 μm^2
- Densité de cellules = 234 cellules par μm^2.

Il n'y a pas de relation entre le nombre de particules en 3D et le nombre de profils de particules vus en section.

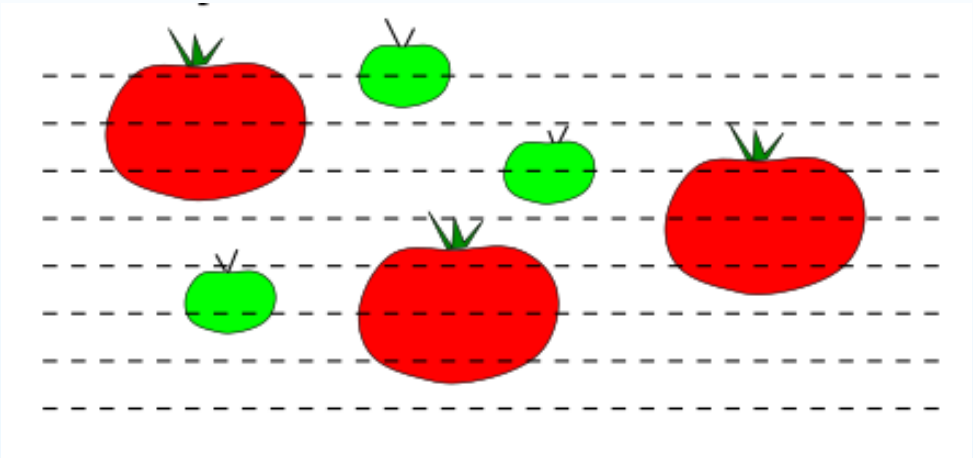
La salade de tomate

Imaginons faire une salade avec 3 grandes tomates rouges et 3 petites vertes.



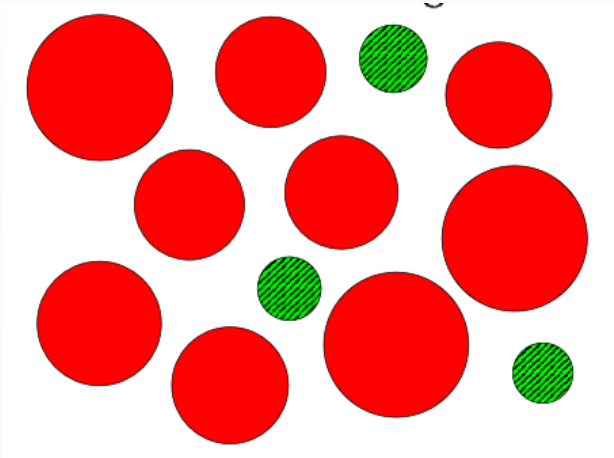
La salade de tomate (II)

On couple la salade régulièrement :



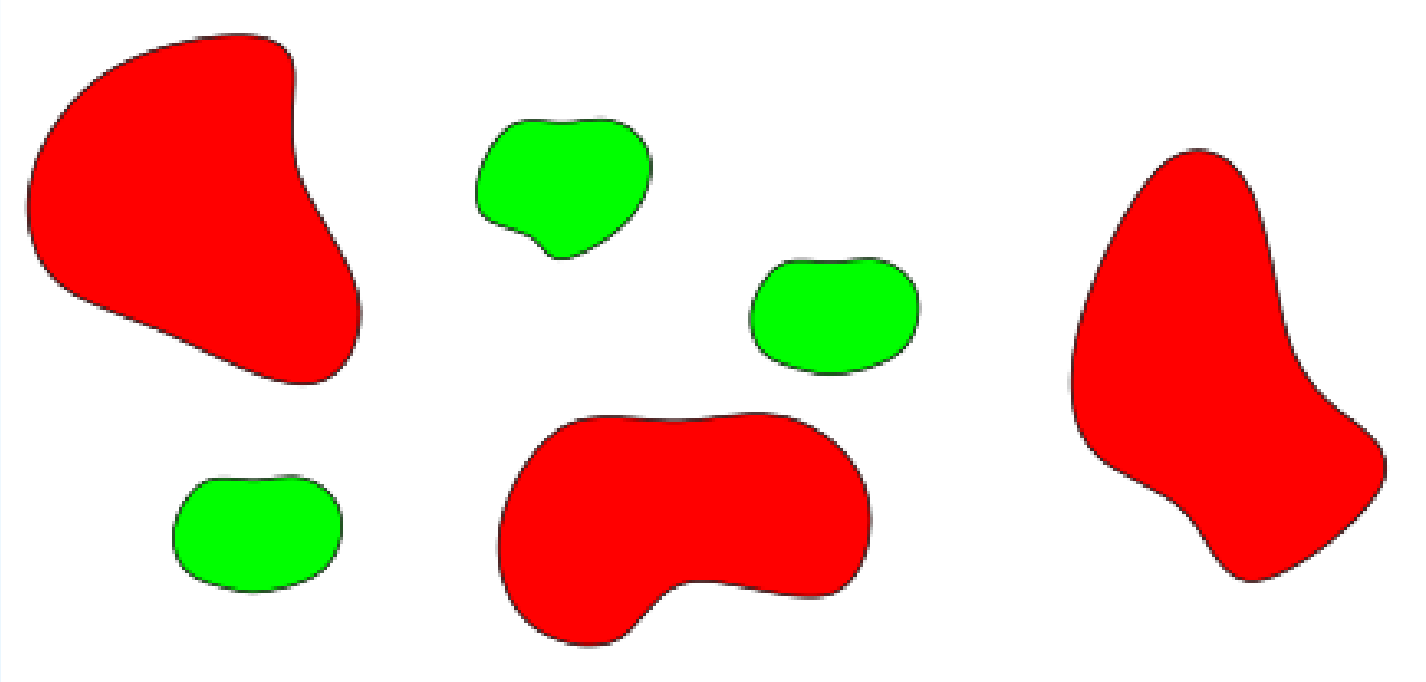
La salade de tomate (III)

On n'a pas 50% de coupes rouges et 50% de coupes vertes !



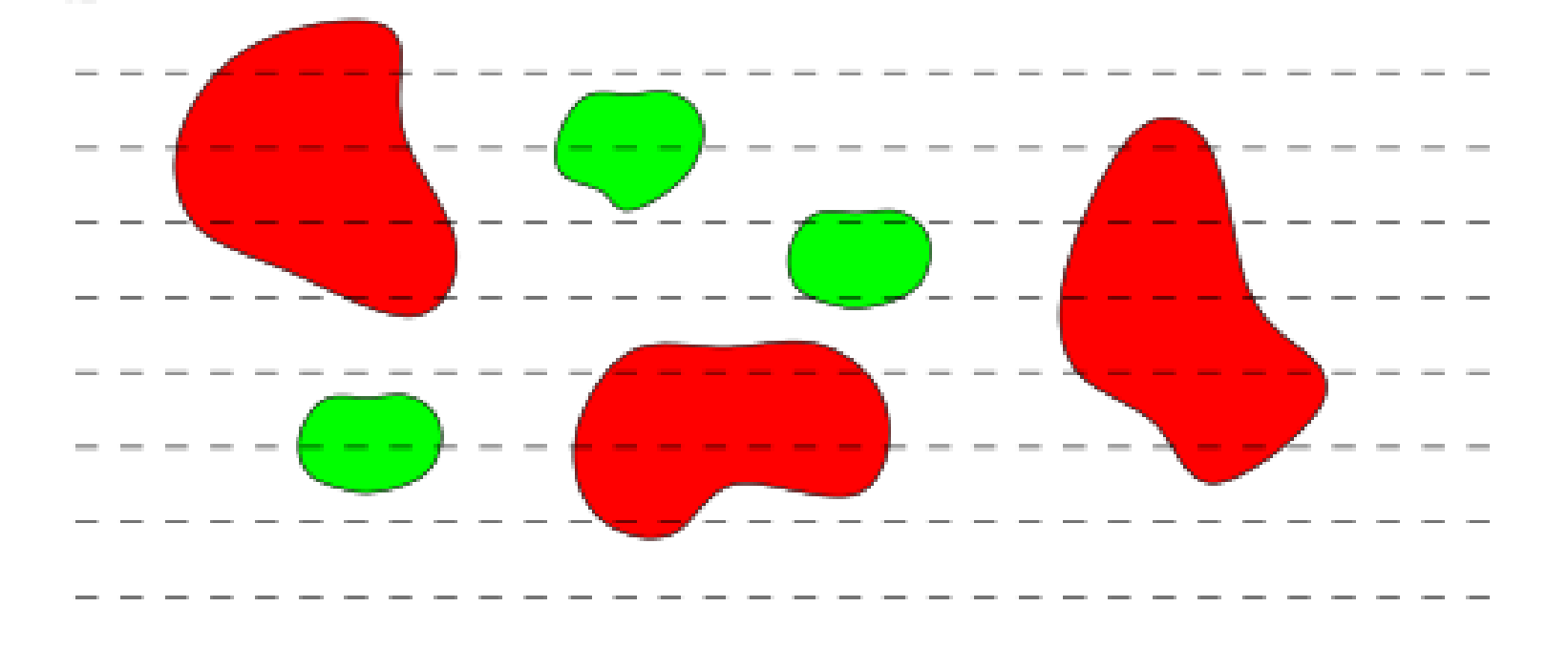
Une population de cellules

Soit une population de cellules



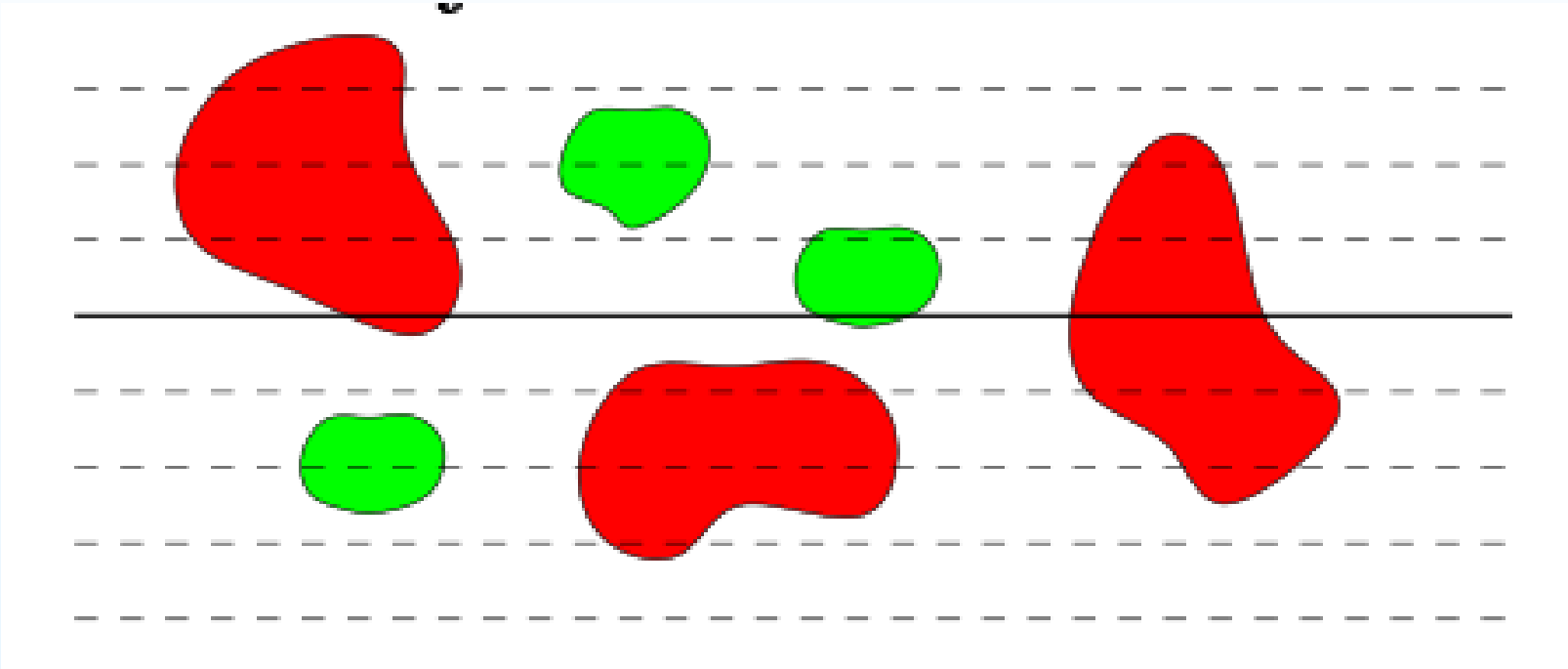
On prend des coupes séries

Comme avant, réparties régulièrement



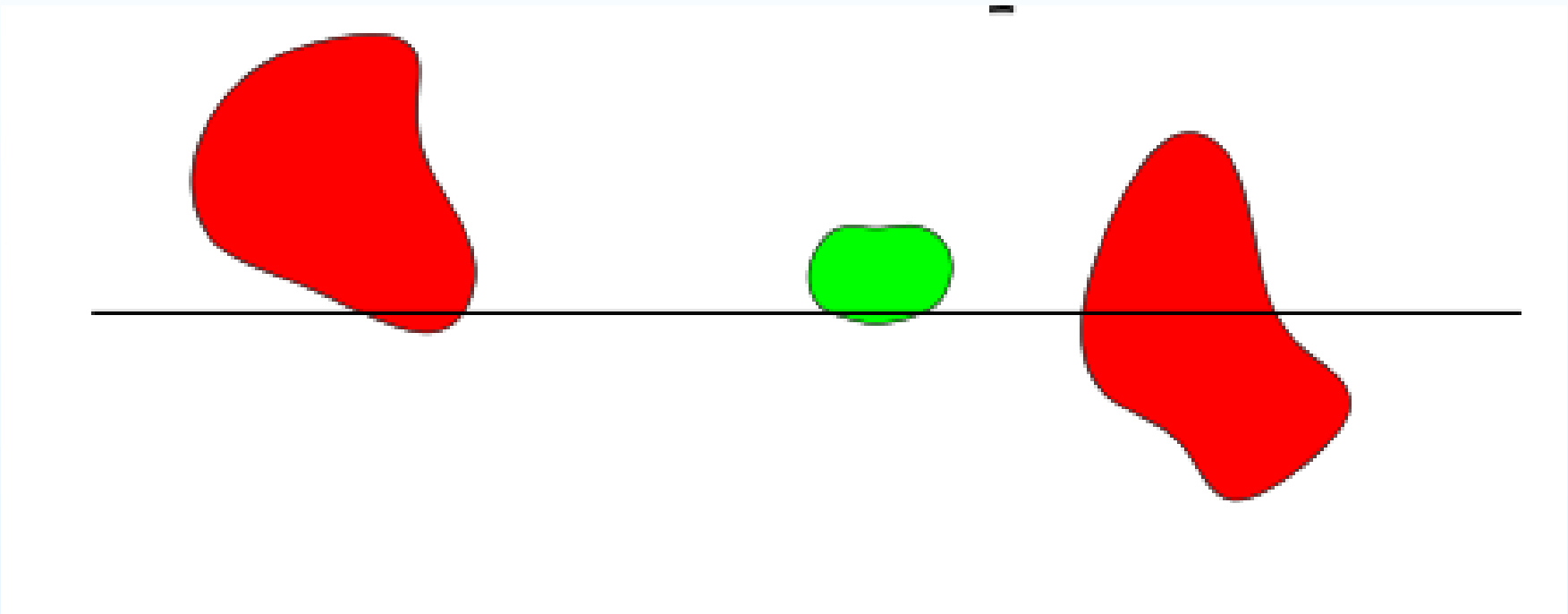
Echantillonnage

On prend une coupe au hasard



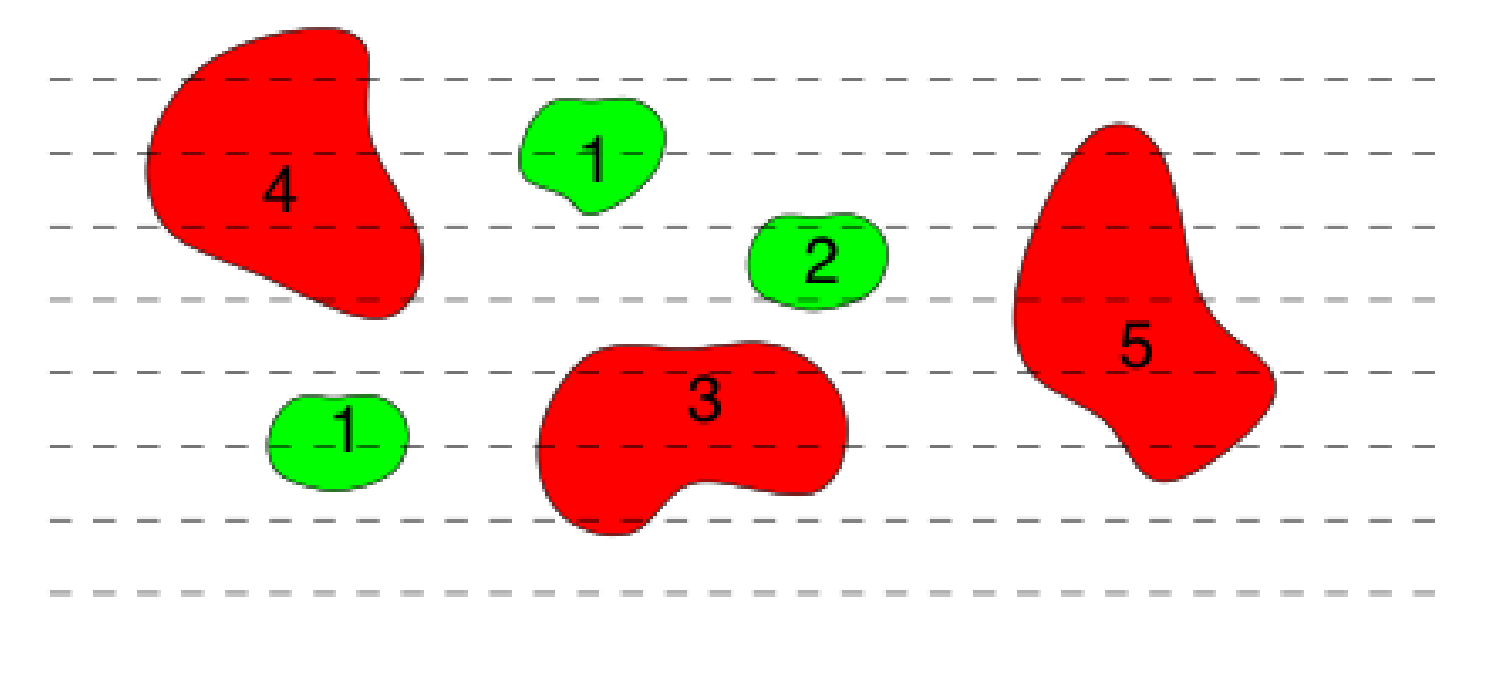
Choix

Voici les cellules échantillonnées



Biais d'échantillonnage

Une section est un échantillon *biaisé* de la population de particules/cellules/grains...



Les particules hautes sont sur-représentées.

Illustration

Nombre de particules

Le nombre de particules dans une section dépend de deux facteurs:

- Le nombre de particules
- La taille (la hauteur) de ces particules

Il n'y a pas de moyen facile d'inverser cette relation en général.

Exceptions

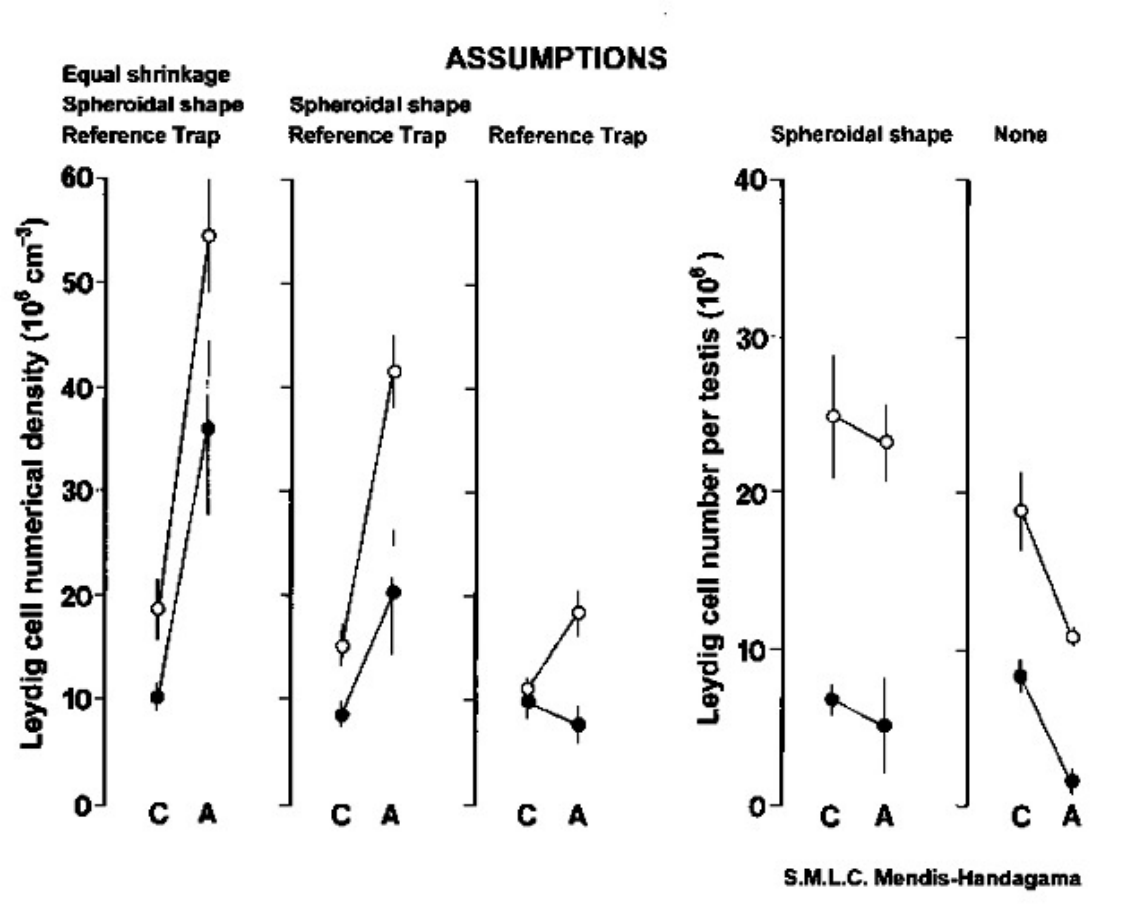
Rhines-DeHoff relation:

$$N_a = \bar{H} N_v$$

Où
N_v = nombre de particules par volume unitaire
N_a = nombre de profils de particules par unité de surface
$\bar{H}$ = hauteur moyenne des particules.
Particules doivent être convexes et homogènes spaciallement.

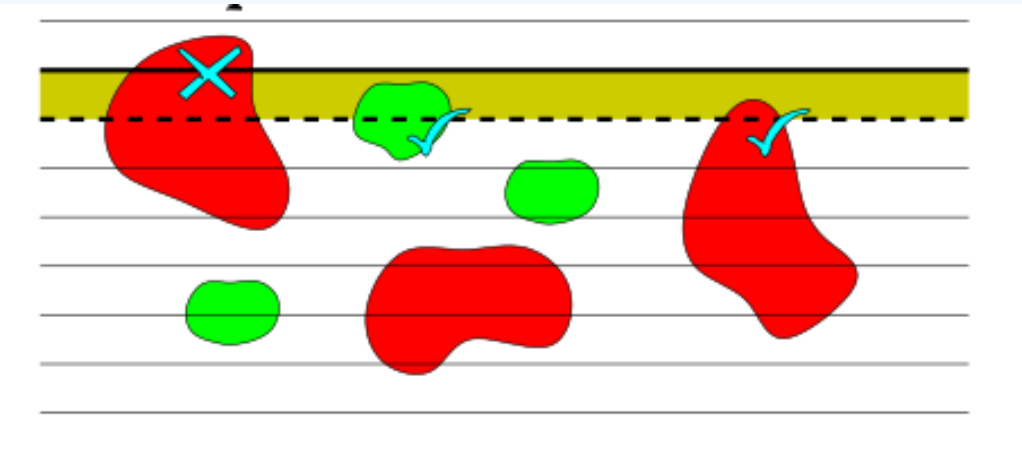
Comptage de neurones

Pour déterminer une densité de neurones, la pratique était de compter la fréquence de neurones dans les sections puis de la “corriger”.



La règle du disecteur

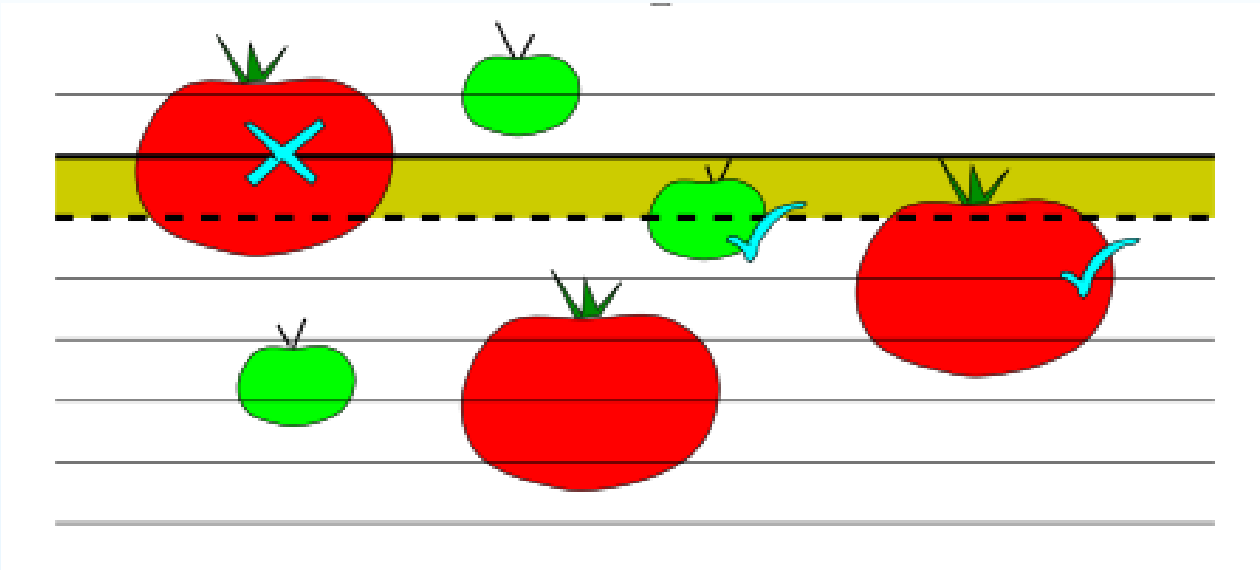
1. Diviser l'espace en section épaisses
2. Prendre un échantillonnage uniforme de sections avec proportion connue f
3. Pour chaque section compter uniquement les particules dont la partie supérieure est contenue dans la section d'intérêt.



4. Les particules choisies forment l'échantillon

Marqueur unique par particule

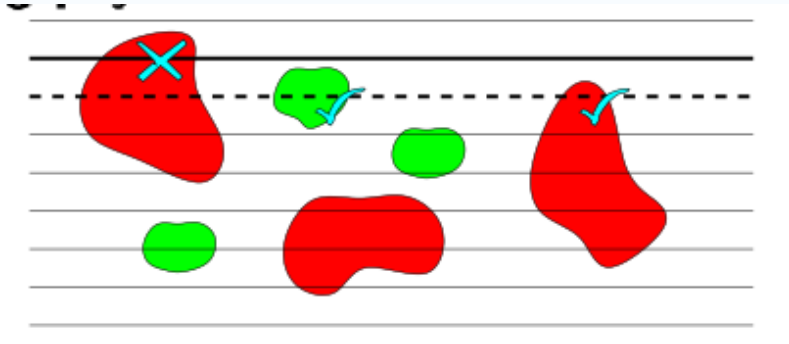
La règle du disecteur permet d'associer un point unique par objet, et donc de l'identifier (le haut de l'objet).



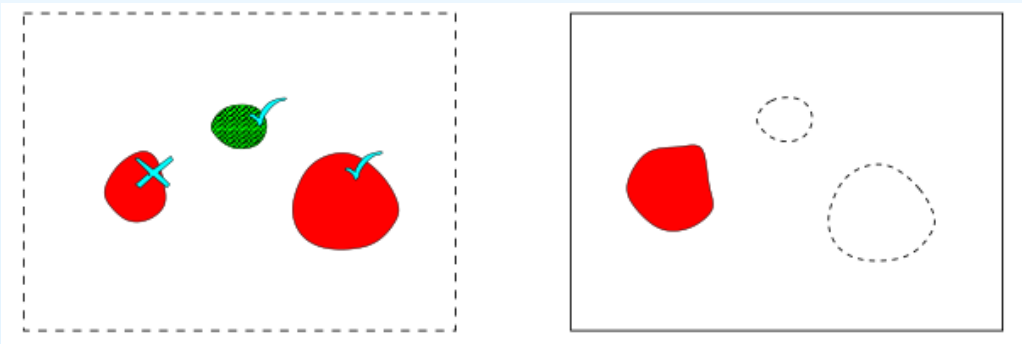
Clairement on obtient un échantillonnage uniforme de cette façon.

Avec des sections fines

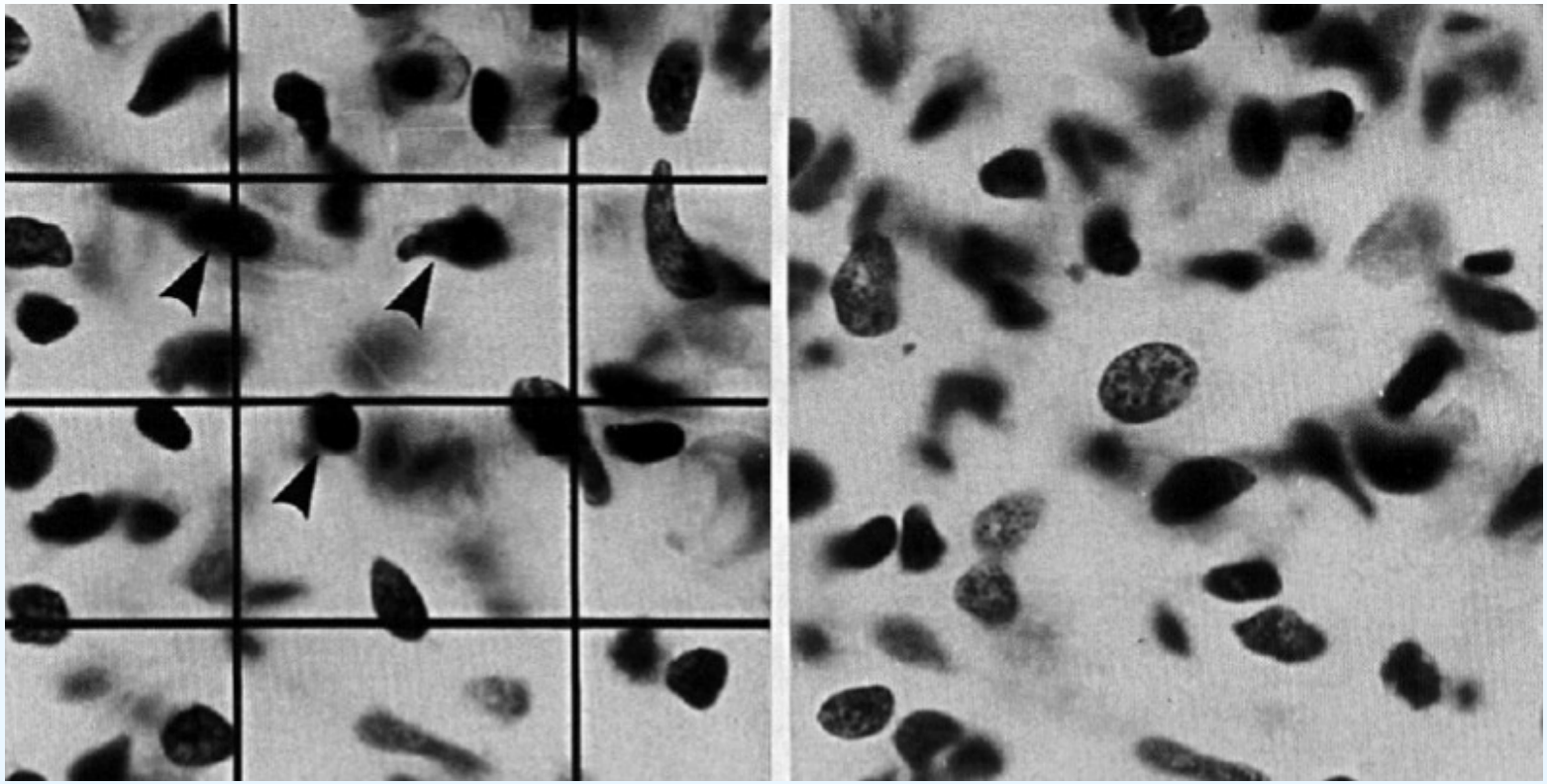
Avec des sections fines, on utilise deux plans.



Une section devient le plan de comptage et l'autre le plan interdit. On ne compte que les sections qui sont dans le plan de comptage et qui n'ont pas de section correspondante dans le plan interdit.



Exemple réel



Invention du disecteur

- W.R. Thompson (1929), dentiste de l'Iowa
- D.R. Cox, H.E. Daniels (1940s), textile
- H. Gundersen (1986), Microscopie

Comment compter les neurones?

1. Découper un cerveau entier en petits morceaux
2. Prendre un échantillon systématique des pièces
3. Couper les pièces en plus petit morceaux encore
4. Prende un échantillon systématique de celles-ci

   ...

n Compter les cellules avec la règle du disecteur. Multiplier par l'inverse de la fraction d'échantillonnage.

Le résultat n'est pas biaisé.

Compter des fibres

1. Que pensez vous de l'approche suivante

Le résultat n'est pas biaisé.

Conclusion

- Certaines techniques standard de microscopie donnent lieu à des biais
- Les principes de base de l'échantillonnage aléatoire rendent possible l'estimation de nombreuses quantités (surface, volume, etc)
- L'échantillonnage systématique est simple et efficace
- On doit utiliser certains protocoles pour éviter les biais de comptage dans le cas du dénombrement d'objets.