

Eléments d'analyse combinatoire - exercices

Probabilités et statistique pour la biologie : éléments (STAT1)

Jacques van Helden

2018-11-26

Exercice 02.1: oligomères sans résidus répétés

1. Combien d'oligomères peut-on former (ADN ou peptides) en utilisant chaque résidu une et une seule fois ?

Approche suggérée: agrégez progressivement les résidus, en vous demandant à chaque étape combien d'entre eux n'ont pas encore été incorporés.

2. Généralisez la formule pour des séquences d'objets tirés dans un ensemble de taille arbitraire (n).
3. Quel est le nom de la fonction donnant le résultat ?
4. Dans ce processus, quel est le mode de sélection des résidus: **avec ou sans remise** ?

Exercice 02.2 : listes (ordonnées) de gènes

On a mené une expérience de transcriptome pour mesurer le niveau d'expression de tous les gènes de la levure. Sachant que le génome comporte 6000 gènes, combien de possibilité existe-t-il pour sélectionner les 15 gènes les plus fortement exprimés (**en tenant compte** de l'ordre relatif de ces 15 gènes) ?

Approche suggérée: comme précédemment, simplifiez le problème en partant de la sélection minimale, et augmentez progressivement le nombre de gènes (1 gène, 2 gènes, ...).

Questions subsidiaires:

- ▶ Trouvez un exemple familier de jeu de pari apparenté à ce problème.
- ▶ Généralisez la formule pour la sélection d'une liste de x gènes dans un génome qui en comporte n .

Exercice 02.3 : ensembles (non-ordonnés) de gènes

Lors d'une expérience de transcriptome indiquant le niveau d'expression de tous les gènes de la levure. Sachant que le génome comporte 6000 gènes, combien de possibilité existe-t-il pour sélectionner les 15 gènes les plus fortement exprimés (**sans tenir compte** de l'ordre relatif de ces 15 gènes) ?

Approche suggérée: comme précédemment, simplifiez le problème en partant de sélections minimales (1 gène, 2 gènes, ...) et généralisez la formule.

Questions subsidiaires:

- ▶ Trouvez un exemple familier de jeu de pari apparenté à ce problème.
- ▶ Généralisez la formule pour la sélection d'un ensemble de x gènes dans un génome qui en comporte n .
- ▶ Connaissez-vous le nom de la formule ainsi trouvée ?

Exercice 02.4 : dénombrement des mutations possibles

On soumet une librairie de molécules d'ADN de 1 kilobase à un traitement mutagène qui provoque un nombre moyen de 5 mutations ponctuelles (substitutions) par molécule. Combien de séquences distinctes existe-t-il pour un fragment muté en 5 positions indépendantes ?

Expliquez comment vous modélisez ce problème. Écrivez les formules avec les symboles, puis remplacez les symboles par les valeurs numériques correspondantes. Il n'est pas nécessaire de calculer la valeur finale (ceci nécessite un ordinateur, nous le ferons pendant les TP).

Exercice 02.5 : oligopeptides 3×20

Combien d'oligopeptides de taille 60 peut-on former en utilisant exactement 3 fois chaque acide aminé ?