**Recon** : une méthode modifiée du maximum de vraisemblance (maximum likelihood) qui calcul la diversité globale d'un répertoire à partir de mesures sur un échantillon. Recon produit des estimations précises et robustes par n'importe quelle mesure de diversité complémentaires, y compris la richesse en espèces et l'entropie, à une couverture de répertoire fractionnaire. Il produit également des barres d'erreur et des tableaux de puissance, permettant des comparaisons robustes de la diversité entre les individus et au fil du temps.

recon.py requires:

Python 2.7 (https://www.python.org)

the SciPy Python library (http://scipy.org)

Plotting additionally requires:

d3.js (http://d3js.org)

wkhtmltopdf (http://wkhtmltopdf.org)

cpdf (http://community.coherentpdf.com)

Recon est un algorithme permettant de générer une description d'une population globale à partir d'un échantillon

Pas de supposition sur la forme de la distribution des espèces dans la population (par exemple, exponentielle, puissance, etc.), pas de Overffiting sur le bruit de l'échantillon, pas de minimum local, et produit des intervalles de confiance à 95 % pour le nombre d'espèces d'une taille donnée dans une population.

Pour le faire marcher j’avais besoin d’installer linux sur VM car sur ubuntu de windows j’avais plein d’erreurs. Après installation j’ai créé un nouvel environnement virtuel pour installer python 2.7 (il ne marche pas sous le python3 j’ai essayé).

Le code fourni :

1. Détermine la distribution de fréquence des espèces de la population globale en utilisant une approche de maximum-likelihood modifiée qui s'arrête avant le Overfitting.
2. Donne des mesures de la diversité pour l'ensemble de la population
3. Génère des barres d'erreur
4. Effectue des calculs de puissance « power calculations » de la taille d'échantillon requise pour pouvoir détecter une différence donnée entre deux populations globales

Je suis bloqué sur l’étape 1. Pourquoi ? Voila ce qui est décris dans le github comme exemple d’application :

python recon.py -R -t 30 -c -o test\_sample\_1\_fitfile.txt test\_sample\_1.txt

Required parameters:

OUTPUT\_FILE (preceded by -o): the filename to be used for output. Note that if this file exists it will not be overwritten; instead, recon will exit with an error message.

INPUT\_FILE: a text file containing the number of individuals of each species seen in the sample (i.e., the species sizes in the sample).

.

.

.

Etc.

Voici ma commande pour tester :

python recon.py -R -t 30 -c -o Raouf\_sample\_1\_fitfile.txt test\_sample\_1.txt

et voici l’erreur :

recon.py: error: unrecognized arguments: -R -o

Malgré que le -R et le -o sont obligatoire (je pense) pour faire marcher la commande.

J’ai cherché cette erreur dans leurs ‘‘issues’’ mais je n’ai pas réussi à trouver la solution.

En tout cas de ce que j’ai compris pour la première application :

Le fichier d’entrée est un fichier .txt qui contient soit :

1. La taille de l’espèce et la nombre d’espèce qui ont cette taille, par exemple

1 1833459

2 405423

3 86822

4 18467

5 3694

6 626

7 128

8 20

9 1

Ici il y a 1833459 espèces qui ont la taille 1 et 9 espèce qui ont la taille 9.

1. Le nom de l’espèce et le nombre d’individu dans cette espèce :

9\_0 9

8\_0 8

8\_1 8

8\_2 8

8\_3 8

8\_4 8

8\_5 8

L’espèce 8\_0 a 8 individus

La sortie (Output) copie-coller de leurs github

([0.88205137420509439, 0.11794862579490561], [0.306026605847127, 1.0723167141789034], 5224621, {1: 1833459, 2: 405423, 3: 86822, 4: 18467, 5: 3694, 6: 626, 7: 128, 8: 20, 9: 1}, None, -1572523.5668443954, 2.2577288150787354)

J’ai besoin de regarder en détail cette partie dés que je puisse faire marcher ma commande.

+ les paramètres optionnels

Le/Les fitfile.txt qui sort de cette première application va Servier pout la deuxième application : le calcul de la diversité

python recon\_v2.1.py -D -Q 0 1 inf -b error\_bar\_parameters.txt -o test\_D\_number\_table.txt test\_sample\_1\_fitfile.txt test\_sample\_2\_fitfile.txt test\_sample\_3\_fitfile.txt

Je regarderai en détail cette partie ainsi que la suite dés que je finisse avec la première partie.