Partiel de Probabilités

1 Jeux de pile/face non transitif

Problème 1 : Motivation biologique : Pour détecter les séquences génétiques sous sélection (donc sûrement fonctionnelles), des méthodes statistiques se sont intéressées à étudier leurs fréquences et les relation entre elles.

Pour simplifier l'analyse, nous allons considérer une séquence de pile ou face symétrique : au n^e tirage, on tire P ou F avec probabilité 0,5 chacun et les tirages sont indépendants. On va voir que les notions de " fréquence d'apparition ", de " temps moyen d'attente " et de " dominance " ne sont pas aussi identiques qu'on pourrait le penser.

1) Quelle est la probabilité d'avoir la séquence PFPF à une position donnée ?

2) En utilisant la linéarité de l'espérance, calculer le nombre moyen d'occurrences de cette séquence PFFP parmi 203 tirages de pile/face.

Indication : on pensera à décomposer ce nombre d'occurrences selon la position où celle-ci pourrait se produire.

3) En déduire que toutes les séquences de même longueur (ici 4 pour PFPF) ont la même fréquence d'apparitions.

On cherche ici à comparer le nombre moyen de tirages nécessaires entre deux séquences successives de PF versus celui entre deux séquences de FF et de l'autre FF. On les notera respectivement T_1 et T_2, et on admettra que ces quantités sont finies.

4) Notons $T_1(F)$ le nombre moyen de tirages supplémentaires avant d'atteindre PF sachant que le premier tirage est face. De même pour $T_1(P)$ où le premier tirage est pile. Avec la loi du deuxième tirage, trouver deux relation entre ces deux quantités.

5) En déduire l'expression de T_1.

6) Faire de même pour T_2. Obtient-on la même valeur ? Comment l'expliquer vu 3) ?

On veut maintenant calculer la probabilité de trouver PF avant FF. On notera p_1 la probabilité de voir gagner PF en commençant par pile et p_2 en commençant par face.

7) De même que pour la question 4, trouver des relations entre ces deux quantités.

8) En déduire la probabilité demandée. Est-elle cohérente avec le résultat en 6) ?

2 Enjeu écologique du taux de croissance

Problème 2 : Michel et Béatrice se disputent sur la définition naturelle du taux de croissance d'une population. Dans leurs modèles élémentaires, ils décrivent l'évolution au cours des générations de la taille de population de poissons. Notons N_n cette taille à la génération n. Celle la génération suivante est donnée par $N_{n+1} = X_{n+1} \times N_n$, où le facteur multiplicatif X_{n+1} reflète l'aléa de l'environnement.
1) Ecrire N_n en fonction des X_i et de N_0. Dans le cas où les X_i sont constants à x, donner la valeur du taux de croissance r tel que N_n évolue comme e^{rn}.

2) On veut dire qu'une espèce est en danger si son taux de croissance est négatif. A quoi cela correspond-il à la question précédente ?

3) Michel défend l'idée de choisir $r_M := \log \mathbb{E}(X)$ comme définition du taux de croissance, tandis que Béatrice juge plus raisonnable de $r_B := \mathbb{E} \log(X)$. Par une inégalité classique, montrer qu'il y a plus d'espèces en danger selon la définition de Béatrice.

4) Calculer l'espérance de N_n. Est-ce que ce résultat va dans le sens de Michel ou de Béatrice ?

5) On suppose que X vaut 0.8 avec probabilité 0.4 et 1.1 avec probabilité 0.6. Que valent r_M, r_B ? Les conclusions sont-elles identiques quant au risque d'extinction ?

6) Relier $\log(N_n/N_0)$ à une loi binomiale dont on précisera les paramètres.

7) La loi des grands nombres donne-t-elle une justification à une notion de taux de croissance ? Qu'est-ce que cela traduit, notamment vis-à-vis de la question 4 ?

8) Rappeler le Théorème de Moivre-Laplace sur les lois binomiales. Peut-on bien l'appliquer à ce cadre ?

9) Avec quelle probabilité la taille de population a-t-elle été divisée par plus de 10 au cours des 50 premières générations ?

10) Donner une approximation à densité pour la loi de N_n pour de grandes valeurs de N.

11) On considère maintenant le cas d'événements catastrophiques, mais rares. On suppose que X vaut 0.01 avec probabilité 0.01 et 1.03 sinon. Que valent r_M, r_B ? Qu'en dîtes-vous ?
Indication : $log(100) \approx 4.61$, $log(1.03) \approx 2.9610^{-2}$.

12) On suppose que N_0 vaut un million et que l'espèce est en grand danger si sa population tombe en-dessous de 100. On admettra le modèle ci-dessus pour l'évolution de N_n tant que ce seuil n'est pas atteint. Tracer quelques réalisations "typiques" de telles dynamiques aléatoires sur 200 générations.

13) Evaluer la probabilité que l'espèce devienne en grand danger lors des 200 premières générations.
Indication : $e^{-2} \approx 0.135$.
14) Quel est plus généralement l'état de la population après 200 générations ?

15) Que concluez-vous vis-à-vis de ces deux définitions r_M et r_B ?

Remarque : Une version assez similaire de ces résultats est parue dans un Pour la Science récent, dans un article consacré à la concentration de richesse. Je vous invite à trouver ce paradoxe du "vide-grenier" dans l'article consacré du numéro 507.